

A representation for exchangeable coalescent trees and generalized tree-valued Fleming-Viot processes

Stephan Gufler*

October 6, 2016

Abstract

We show that every exchangeable random semi-ultrametric on \mathbb{N} can be obtained by sampling an iid sequence from a random marked metric measure space and adding the marks to the distances. We use this representation to define tree-valued Fleming-Viot processes from the Ξ -lookdown model. The case with dust is included for processes with values in the space of marked metric measure spaces, and for processes with values in the space of distance matrix distributions.

Keywords: Ultrametric, jointly exchangeable array, marked metric measure space, dust, tree-valued Fleming-Viot process, lookdown model, Ξ -coalescent.

AMS MSC 2010: Primary 60G09, Secondary 60J25, 60K35, 92D10

Contents

1	Introduction	2
1.1	Some background on coalescent trees, ultrametrics, and metric measure spaces	2
1.2	The sampling representation	4
1.3	Evolving genealogies	5
1.4	Additional related literature	6
2	Distance matrices and their decompositions	7
3	Sampling from marked metric measure spaces	8
3.1	Preliminaries	8
3.2	Tree-like marked metric measure spaces	8
3.3	Dust-free semi-ultrametrics and dust-free marked metric measure spaces	11
3.4	Marked metric measure spaces from marked distance matrices	11
3.5	The sampling representation	13
4	Application to tree-valued processes	16
5	Genealogy in the lookdown model	19
6	The Ξ-lookdown model	21
6.1	Properties of the genealogy at a fixed time	22
7	Decomposition of the genealogical distances	24
7.1	The deterministic construction	24
7.2	Stochastic evolution	26

*Goethe-Universität, Institut für Mathematik, Postfach 11 19 32, Fach 187, 60054 Frankfurt am Main, Germany, gufler@math.uni-frankfurt.de

8	Tree-valued Fleming-Viot processes	29
8.1	Processes with values in the space of metric measure spaces	29
8.2	Processes with values in the space of marked metric measure spaces	30
8.3	Processes with values in the space of distance matrix distributions	30
9	Some semigroup properties	31
10	Convergence to equilibrium	33
11	Proofs	34
11.1	Measurability of the construction of (marked) metric measure spaces	34
11.2	Construction of the marked metric measure space in the sampling representation	35
12	Construction from the flow of bridges	41

1 Introduction

1.1 Some background on coalescent trees, ultrametrics, and metric measure spaces

In population genetics, coalescents are common models for the genealogy of a sample from a population. The Kingman coalescent [31] is a partition-valued process in which each individual of the sample forms its own block at time 0, and as we look into the past, each pair of blocks merges independently at constant rate. These blocks stand for the families of individuals that have a common ancestor at given times in the past. Generalizations of the Kingman coalescent include the Λ -coalescent (Pitman [41], Sagitov [43], Donnelly and Kurtz [16]) where multiple blocks are allowed to merge to a single block at the same time, and the Ξ -coalescent (Möhle and Sagitov [38], Schweinsberg [44]) where several clusters of blocks may also merge simultaneously.

A realization of a coalescent for an infinite sample can be expressed as a càdlàg path $(\pi_t, t \in \mathbb{R}_+)$ with values in the space of partitions of \mathbb{N} such that π_t is a coarsening of π_s for all $s \leq t$. Assuming that each pair of integers is in a common block of π_t for t sufficiently large, $(\pi_t, t \in \mathbb{R}_+)$ can equivalently be expressed as a semi-ultrametric ρ on \mathbb{N} such that for all $t \in \mathbb{R}_+$ and $i, j \in \mathbb{N}$,

$$\rho(i, j) \leq 2t \quad \text{if and only if } i \text{ and } j \text{ are in the same block of } \pi_t, \quad (1.1)$$

and (1.1) yields a one-to-one correspondence between these càdlàg paths and the semi-ultrametrics on \mathbb{N} , cf. [20, Example 3.41] and [21, p. 34]. A (semi-)ultrametric ρ can also be defined as a (semi-)metric that satisfies the strong triangle inequality

$$\max\{\rho(x, y), \rho(y, z)\} \geq \rho(x, z).$$

Evans [19] studies the completion of the random ultrametric space associated with the Kingman coalescent which he endows with a probability measure such that the mass on each ball is given by the asymptotic frequency of the corresponding family, and a class of more general coalescents is studied by Berestycki et al. [3].

Remark 1.1. Let us briefly recall the well-known correspondence between ultrametric spaces and real trees to which we will refer to explain main concepts in this article. A real tree is a metric space (T, d) that is tree-like in the sense that (i) no subspace is homeomorphic to the unit circle, and (ii) for each $x, y \in T$, there exists an isometry ι

from the real interval $[0, d(x, y)]$ to T with $\iota(0) = x$ and $\iota(d(x, y)) = y$, see e.g. Evans [20] for an overview. An ultrametric space (X, ρ) can be isometrically embedded into the real tree (T, d) that is obtained by identifying the elements with distance zero of the semi-metric space $(\mathbb{R}_+ \times X, d)$ given by $d((s, i), (t, j)) = \max\{\rho(i, j) - s - t, 0\}$, cf. e.g. [21, p. 34]. Clearly, (X, ρ) is isometric to the subspace $\{0\} \times X$ of the leaves of (T, d) . For a semi-ultrametric space (X, ρ) , we identify the elements with distance zero to obtain an ultrametric space which we associate with a real tree (T, d) as above.

As in Remark 1.1, a semi-ultrametric on \mathbb{N} can be considered as an infinite tree whose leaves are labeled by the elements of \mathbb{N} . Often these labels are not relevant, for instance, when they only record the order in which iid samples from a population are drawn. To remove the labels, we could pass to isometry classes. However, the asymptotic block frequencies in the coalescent given by an ultrametric on \mathbb{N} are not determined by the isometry class, as one may apply an infinite permutation without changing the isometry class. To retain just this information besides the metric structure, we can take a measure-preserving isometry class of the completion of the ultrametric space that is endowed with a probability measure that charges each ball with the asymptotic frequency of the corresponding block, if such a probability measure exists. This probability measure can equivalently be described as the weak limit of the uniform probability measures on the individuals $1, \dots, n$, as $n \rightarrow \infty$. Then we obtain the description by isomorphy classes of metric measure spaces of Greven, Pfaffelhuber, and Winter [24] that was applied to Λ -coalescents in the dust-free case. We speak of the dust-free case if the semi-ultrametric space has no isolated points, which means that the coalescent tree has no isolated leaves. Greven, Pfaffelhuber, and Winter [24] also show that their approach is not directly applicable to Λ -coalescents with dust. The most elementary example for the case with dust is the star-shaped coalescent which starts in the partition into singleton blocks which all merge into a single block at some instant. The associated ultrametric on \mathbb{N} induces the discrete topology. Here the uniform probability measures on $1, \dots, n$ do not converge weakly as they converge vaguely to the zero measure.

A metric measure space is a triple (X, r, μ) that consists of a complete and separable metric space (X, r) and a probability measure μ on the Borel sigma algebra on X . An important feature is that one can consider the matrix $(r(x(i), x(j)))_{i, j \in \mathbb{N}}$ of the distances between μ -iid samples $(x(i))_{i \in \mathbb{N}}$. The distribution of $(r(x(i), x(j)))_{i, j \in \mathbb{N}}$ is called the distance matrix distribution of (X, r, μ) . By the Gromov reconstruction theorem (see Theorem 4 of Vershik [47]), there exists a measure-preserving isometry between the supports of the measures of any two metric measure spaces that have the same distance matrix distribution, in which case we call them isomorphic. Under an appropriate condition, Theorem 5 of [47] associates with any typical realization of a random semi-metric on \mathbb{N} a metric measure space whose distance matrix distribution is the distribution of this semi-metric. In the ultrametric case, this metric measure space is the completion of a typical realization of the semi-metric, endowed with the probability measure given by the asymptotic block frequencies of the associated coalescent (as in Remark 3.9). This can also be deduced from [47, Equation (9)].

1.2 The sampling representation

We view a random semi-metric ρ on \mathbb{N} as the random matrix $(\rho(i, j))_{i, j \in \mathbb{N}}$, and we call it exchangeable if $(\rho(i, j))_{i, j \in \mathbb{N}}$ is distributed as $(\rho(p(i), p(j)))_{i, j \in \mathbb{N}}$ for each (finite) permutation p of \mathbb{N} . We give a representation for all exchangeable random semi-ultrametrics on \mathbb{N} in terms of sampling from random marked metric measure spaces. Marked metric measure spaces are introduced in Depperschmidt, Greven, and Pfaffelhuber [12]. A (\mathbb{R}_+) -marked metric measure space is a triple (X, r, m) that consists of a complete and separable metric space (X, r) and a probability measure m on the Borel sigma algebra on the product space $X \times \mathbb{R}_+$. The marked distance matrix distribution of a marked metric measure space (X, r, m) is defined as the distribution of $((r(x(i), x(j))))_{i, j \in \mathbb{N}}, (v(i))_{i \in \mathbb{N}}$ where $(x(i), v(i))_{i \in \mathbb{N}}$ is an m -iid sequence in $X \times \mathbb{R}_+$. We call two marked metric measure spaces isomorphic if they have the same marked distance matrix distribution. We endow the space of isomorphy classes of marked metric measure spaces with the Gromov-weak topology in which marked metric measure spaces converge if and only if their marked distance matrix distributions converge weakly. This yields a Polish space, as shown in [12].

In the present article, we use marked metric measure spaces to obtain from a random variable (r, v) that has the marked distance matrix distribution of a marked metric measure space an exchangeable semi-metric ρ on \mathbb{N} by

$$\rho(i, j) = (r(i, j) + v(i) + v(j)) \mathbf{1}\{i \neq j\}.$$

We call the distribution of $(\rho(i, j))_{i, j \in \mathbb{N}}$ the distance matrix distribution of the marked metric measure space. The basic result in this article is that every exchangeable semi-ultrametric on \mathbb{N} can be represented as the outcome of a two-step random experiment, where we have a random marked metric measure space in the first step, and we sample from this marked metric measure space according to the distance matrix distribution in the second step.

Theorem 1.2. *Let ρ be an exchangeable semi-ultrametric on \mathbb{N} . Then there exists a random variable χ with values in the space of isomorphy classes of marked metric measure spaces such that the law of ρ equals $\mathbb{E}[\nu]$, where ν denotes the distance matrix distribution of χ . That is, for a random semi-metric ρ' on \mathbb{N} with conditional distribution ν given χ , the random variables ρ and ρ' have the same (unconditional) distribution. The random variable χ is unique in distribution.*

A stronger formulation of Theorem 1.2 is given in Theorem 3.18 where we state the marked metric measure space χ realizationwise as a deterministic function of ρ' . The key idea for the construction of χ is to decompose the tree that is associated with ρ' into the external branches and the remaining subtree. Here we define that an external branch consists only of the leaf if that leaf corresponds to an integer that has ρ' -distance zero to another integer. In the marked metric measure space χ , the marks encode the external branch lengths, and the metric space describes the remaining subtree. The external branches all have length zero in particular in the dust-free case. In this case, χ can also be replaced by the isomorphy class of a metric measure space (as in Section 3.3).

We prove the main part of Theorem 3.18 in Section 11.2. We formulate the decomposition at the external branches in Section 2 in terms of semi-ultrametrics. As we discuss

in Section 3.2 and in Remark 3.15, marked metric measure spaces whose distance matrix distribution is concentrated on semi-ultrametrics correspond to real trees that are endowed with a probability measure that is concentrated on the starting vertices of the external branches. We mention that Evans, Grübel, and Wakolbinger [21] also decompose real trees into the external branches and the remaining subtree to give a representation of the elements of the Doob-Martin boundary of Rémy's algorithm in terms of sampling from a weighted real tree and an additional structure. In [21, Section 7], a sampling representation for exchangeable ultrametrics is considered (see Remark 11.2).

We define the action of a finite permutation p of \mathbb{N} on a semi-ultrametric ρ on \mathbb{N} by $p(\rho) = (\rho(p(i), p(j)))_{i, j \in \mathbb{N}}$. The distance matrix distribution of a marked metric measure space (X, r, m) is invariant and ergodic with respect to the action of the group of finite permutations. Indeed (analogously to [47, Lemma 7]), a finite permutation of the distance matrix distribution corresponds to a finite permutation of the m -iid sequence $(x(i), v(i))_{i \in \mathbb{N}}$ which is invariant and ergodic with respect to finite permutations, and from which the distance matrix distribution is obtained as the distribution of $((v(i) + r(x(i), x(j)) + v(j)) \mathbf{1}\{i \neq j\})$. Hence, Theorem 1.2 decomposes the distribution of an exchangeable semi-ultrametric ρ into ergodic components. The ergodic component, that is, the distance matrix distribution of χ , can be read off from each typical realization of ρ by Theorem 3.18 (see also Remark 3.23). The ergodic component is also characterized by the marked metric measure space χ itself. In the dust-free case, it can also be expressed as the isomorphy class of a metric measure space. The finite analog of the aforementioned ergodic decomposition is that a (discrete) random tree whose leaves are labeled exchangeably can be obtained by first drawing the random unlabeled tree and then sampling the labels of the leaves uniformly without replacement. The representation given by Theorem 1.2 can also be seen in the context of the more general but less explicit Aldous-Hoover-Kallenberg representation, see e. g. [29, Section 7].

1.3 Evolving genealogies

In Section 4, we lay the foundation for our study of evolving genealogies by considering a Markov process with values in the space of semi-ultrametrics on \mathbb{N} ; this process describes evolving leaf-labeled trees. Assuming that the state at each time is exchangeable, we map this process to the processes of the ergodic components, expressed as isomorphy classes of metric measure spaces, marked metric measure spaces, and distance matrix distributions, respectively, as outlined in Subsection 1.2. Using the criterion of Rogers and Pitman [42, Theorem 2], we deduce that these image processes are also Markovian, and we describe them by well-posed martingale problems. This is an example of Markov mapping in the sense of Kurtz [33], and Kurtz and Nappo [34].

In Sections 5 – 7, we study a concrete Markov process with values in the space of semi-ultrametrics, namely the process given by the evolving genealogical trees in a lookdown model with simultaneous multiple reproduction events. Lookdown models were introduced by Donnelly and Kurtz [15, 16] to represent measure-valued processes along with their genealogy, see also e. g. Etheridge and Kurtz [17] and Birkner et al. [8]. A lookdown model can be seen as a (possibly) infinite population model in which each individual at each time is assigned a level. The role of this level is model-inherent, namely to order the individuals such that the restriction of the model to the first finitely many

levels is well-behaved (i.e. only finitely many reproduction events are visible in bounded time intervals) and that the modeled quantity (e.g. types, genealogical distances) remains exchangeable. In [16] and in the present article, the level is the rank among the individuals at the respective time according to the time of the latest descendant. Although the levels in finite restrictions of the lookdown model differ from the labels in the Moran model, the processes of the unlabeled genealogical trees coincide which is used to study the length of the genealogical trees in Pfaffelhuber, Wakolbinger, and Weisshaupt [40] and Dahmer, Knobloch, and Wakolbinger [11].

In Section 8, we remove the labels from the process of the evolving genealogical trees in the infinite lookdown model by applying the result from Section 4 to the process from Sections 5 – 7. We call the processes of the ergodic components tree-valued Fleming-Viot processes, regardless which one of the three state spaces we use. The tree-valued Fleming-Viot process with values in the space of isomorphy classes of metric measure spaces is introduced in the case with binary reproduction events (which is associated with the Kingman coalescent) by Greven, Pfaffelhuber, and Winter [25] as the solution of a well-posed martingale problem that is the limit in distribution of corresponding processes read off from finite Moran models. In [25, Remark 2.20], a construction of (a version of) this process from the lookdown model of Donnelly and Kurtz [15] is outlined. The aim in the present article regarding tree-valued Fleming-Viot process is the generalization to the case with dust. We remark that tree-valued Fleming-Viot processes with mutation and selection are studied in Depperschmidt, Greven, and Pfaffelhuber [13, 14] where the states are isomorphy classes of marked metric measure spaces and the marks encode allelic types. In the present article, the marks encode lengths of external branches. We consider only the neutral case, and we describe genealogies without using types.

In Section 9, we show continuity properties of the semigroups of tree-valued Fleming-Viot processes and that the domains of the martingale problems for them are cores. In Section 10, we show that tree-valued Fleming-Viot processes converge in distribution to equilibrium.

1.4 Additional related literature

Aldous [1] represents consistent families of finite trees that satisfy a “leaf-tight” property by random measures on ℓ_1 (and random subsets of ℓ_1). Kingman’s coalescent is given as an example in [1]. The “leaf-tight” property corresponds to the absence of dust. A representation for exchangeable hierarchies in terms of sampling from random weighted real trees is given by Forman, Haulk, and Pitman [22]. There are many other representation results for exchangeable structures in the literature. For instance, by the Dovbysh-Sudakov theorem, see Austin [2] for a proof based on a representation for exchangeable random measures, jointly exchangeable arrays that are non-negative definite can be represented in terms of sampling from the space $L_2[0, 1] \times \mathbb{R}_+$.

The genealogy in the lookdown model is further studied in Pfaffelhuber and Wakolbinger [39]. Kliem and Löhner [32] further study marked metric measure spaces. In their article, tree-valued Λ -Fleming-Viot processes in the dust-free case and an application of marked metric measure spaces to trait-depending branching are also mentioned. In the context of measure-valued spatial Λ -Fleming-Viot processes with dust, Véber and Wakolbinger [46] work with a skeleton structure. Functionals of coalescents like external

branch lengths have also been studied, see for example [37]. Also the time evolution of such functionals has been studied for evolving coalescents, see for example [10, 30].

Bertoin and Le Gall [5–7] represent Ξ -coalescents in terms of sampling from flows of bridges from which they also construct measure-valued Fleming-Viot processes. They also consider mass coalescents. Mass coalescents (see e.g. Chapter 4.3 in Bertoin [4]) also describe genealogies without labeling individuals. In Section 12, we construct the Fleming-Viot process with values in the space of distance matrix distributions from the dual flow of bridges. We also mention the work of Labbé [35] where relations between the lookdown model and flows of bridges are studied.

While we construct versions of tree-valued Fleming-Viot processes in the present article using the representation result (see Remark 8.2), a pathwise construction that uses techniques specific to the lookdown model is given in the companion article [27].

2 Distance matrices and their decompositions

We write $\mathbb{N} = \{1, 2, 3, \dots\}$. Let \mathfrak{U} denote the space of semi-ultrametrics on \mathbb{N} and let \mathfrak{D} denote the space of semimetrics on \mathbb{N} . We view \mathfrak{U} and \mathfrak{D} as subspaces of $\mathbb{R}^{\mathbb{N}^2}$ in that we do not distinguish between a semi-metric ρ and the distance matrix $(\rho(i, j))_{i, j \in \mathbb{N}}$. We endow $\mathbb{R}^{\mathbb{N}^2}$ with a complete and separable metric that induces the product topology when \mathbb{R} is equipped with the Euclidean topology. Using the map

$$\alpha : \mathbb{R}_+^{\mathbb{N}^2} \times \mathbb{R}_+^{\mathbb{N}} \rightarrow \mathbb{R}_+^{\mathbb{N}^2}, \quad (r, v) \mapsto ((v(i) + r(i, j) + v(j)) \mathbf{1}\{i \neq j\})_{i, j \in \mathbb{N}},$$

we define the space

$$\hat{\mathfrak{U}} = \{(r, v) \in \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} : \alpha(r, v) \in \mathfrak{U}\}$$

whose elements we call decomposed semi-ultrametrics or marked distance matrices. As above, we view $\mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$ and $\hat{\mathfrak{U}}$ as subspaces of $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ which we endow with a complete and separable metric that induces the product topology.

We define the function

$$\Upsilon : \mathfrak{U} \rightarrow \mathbb{R}_+^{\mathbb{N}}, \quad \rho \mapsto \left(\frac{1}{2} \inf_{j \in \mathbb{N} \setminus \{i\}} \rho(i, j)\right)_{i \in \mathbb{N}},$$

and we denote by β the function that maps a semi-ultrametric $\rho \in \mathfrak{U}$ to the decomposed semi-ultrametric $(r, v) \in \hat{\mathfrak{U}}$ that is given by $v = \Upsilon(\rho)$ and

$$r(i, j) = (\rho(i, j) - v(i) - v(j)) \mathbf{1}\{i \neq j\}$$

for $i, j \in \mathbb{N}$. The interpretation of these functions is given in Remark 2.2 below from which it follows that r is a tree-like semi-metric (i.e., r is 0-hyperbolic, see e.g. [20]). Alternatively, it can be easily checked that r satisfies the triangle inequality.

The function α retrieves the semi-ultrametric from a decomposed semi-ultrametric. For instance, $\alpha \circ \beta$ is the identity map on \mathfrak{U} .

Remark 2.1. Let us agree on the following notation. When we identify the elements of a semi-metric space (X, ρ) that have ρ -distance zero to obtain a metric space (X', ρ) , we refer by each element $x \in X$ also to the associated element of X' . Furthermore, we define the metric completion of the semi-metric space (X, ρ) as the metric completion of (X', ρ) .

Remark 2.2. Let $\rho \in \mathfrak{U}$, $(r, v) = \beta(\rho)$, and let (T, d) be the real tree associated with ρ as in Remark 1.1 with $X = \mathbb{N}$. Then $v(i) = \Upsilon(\rho)(i)$ can be interpreted as the length, and $(v(i), i)$ as the starting vertex of the external branch that ends in the leaf $(i, 0)$ of T . Here we define that this external branch consists only of the leaf if there exists $k \in \mathbb{N} \setminus \{i\}$ with $\rho(i, k) = 0$. Furthermore, the map $\varphi(i) = (v(i), i)$ from (\mathbb{N}, r) to (T, d) is distance-preserving.

In this sense, the map $\beta : \rho \mapsto (r, v)$ decomposes the coalescent tree that is given by ρ into the external branches with lengths v and the subtree spanned by their starting vertices whose mutual distances are given by r . More generally, any element of $\hat{\mathfrak{U}}$ can be seen as a decomposed coalescent tree.

3 Sampling from marked metric measure spaces

3.1 Preliminaries

Recall the space of isomorphism classes of metric measure spaces from Section 1.1. We denote this space by \mathbb{M} and we endow it with the Gromov-weak topology in which metric measure spaces converge if and only if their distance matrix distributions converge. Greven, Pfaffelhuber, and Winter [24] showed that \mathbb{M} is then a Polish space. Also recall from Section 1.2 the space of isomorphism classes of marked metric measure spaces which we denote by $\hat{\mathbb{M}}$. The distance matrix distribution of the isomorphism class of a metric measure space $\chi \in \mathbb{M}$ is denoted by ν^χ . The marked distance matrix distribution of $\chi' \in \hat{\mathbb{M}}$ is denoted by $\nu^{\chi'}$, so that $\alpha(\nu^{\chi'})$ is the distance matrix distribution of χ' , in accordance with the definition in Section 1.2.

Let S_∞ denote the group of finite permutations on \mathbb{N} . We define the action of S_∞ on \mathfrak{D} and $\mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$, respectively, by $p(\rho) = (\rho(p(i), p(j)))_{i, j \in \mathbb{N}}$ and

$$p(r, v) = ((r(p(i), p(j)))_{i, j \in \mathbb{N}}, (v(p(i)))_{i \in \mathbb{N}})$$

for $p \in S_\infty$, $\rho \in \mathfrak{D}$, $(r, v) \in \mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$. A random variable, for instance with values in \mathfrak{D} or $\mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$, is called exchangeable if its distribution is invariant under the action of the group S_∞ .

Remark 3.1. Exchangeable random variables with values in \mathfrak{D} or $\mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$ can be seen as jointly exchangeable arrays, see e.g. [29, Section 7]. Also recall that the definition of exchangeability does not change when S_∞ is replaced with the group of all bijections from \mathbb{N} to itself, as the finite restrictions determine the distribution of a random variable in \mathfrak{D} or $\mathfrak{D} \times \mathbb{R}_+^\mathbb{N}$.

Remark 3.2. The coalescents associated by (1.1) with the exchangeable semi-ultrametrics on \mathbb{N} form a larger class of processes than the so-called exchangeable coalescents defined in e.g. Section 4.2.2 of Bertoin [4]. For example, the coalescent process associated with an exchangeable semi-ultrametric on \mathbb{N} needs not be Markovian.

3.2 Tree-like marked metric measure spaces

We consider the space

$$\mathbb{U} = \{\chi \in \mathbb{M} : \nu^\chi(\mathfrak{U}) = 1\}$$

of ultrametric measure spaces which is a closed subspace of \mathbb{M} , as shown in [25, Lemma 2.3]. By the same argument, the space

$$\hat{\mathbb{U}} = \{\chi \in \hat{\mathbb{M}} : \alpha(\nu^\chi)(\mathfrak{U}) = 1\}.$$

is a closed subspace of $\hat{\mathbb{M}}$. It contains the marked metric measure spaces with ultrametric distance matrix distribution. Following e. g. [24, 25], we call the elements of \mathbb{U} trees. Also the elements of $\hat{\mathbb{U}}$ may be called trees, as we discuss in this subsection.

First we show in Proposition 3.3 that a. e. realization of a $\hat{\mathfrak{U}}$ -valued random variable with the marked distance matrix distribution of a marked metric measure space in $\hat{\mathbb{U}}$ is the decomposition of a semi-ultrametric by the map β from Section 2. As a consequence, the isomorphy class of a marked metric measure space in $\hat{\mathbb{U}}$ is determined already by its distance matrix distribution.

Proposition 3.3. *Let (X, r', m) be a marked metric measure space with $\alpha(\nu^{(X, r', m)})(\mathfrak{U}) = 1$. Let (r, v) be a $\hat{\mathfrak{U}}$ -valued random variable with distribution $\nu^{(X, r', m)}$. Then*

$$(r, v) = \beta \circ \alpha(r, v) \quad \text{a. s.}$$

The proof relies on the fact that an iid sequence in a separable metric space has no isolated points.

Proof. Let $((x(i), v(i)), i \in \mathbb{N})$ be an m -iid sequence in $X \times \mathbb{R}_+$. We may assume

$$r = (r(i, j))_{i, j \in \mathbb{N}} = (r'(x(i), x(j)))_{i, j \in \mathbb{N}}.$$

We write $\rho = \alpha(r, v)$. We show that $v = \Upsilon(\rho)$ a. s. from which the assertion follows by definition of the map β .

Let $\varepsilon > 0$ and $i \in \mathbb{N}$. By separability, $X \times \mathbb{R}_+$ can be covered by countably many balls of diameter ε . This implies

$$m\{(x', v') \in X \times \mathbb{R}_+ : r'(x(i), x') \vee |v(i) - v'| \leq 2\varepsilon\} > 0 \quad \text{a. s.,}$$

and that there exists a random $j \in \mathbb{N} \setminus \{i\}$ with

$$r'(x(i), x(j)) \vee |v(i) - v(j)| \leq 2\varepsilon \quad \text{a. s.} \tag{3.1}$$

By inequality (3.1) and the definition of ρ , it follows that

$$2v(i) + 4\varepsilon \geq v(i) + v(j) + r(i, j) = \rho(i, j) \quad \text{a. s.}$$

Using the definition of the map Υ , we deduce

$$v(i) + 2\varepsilon \geq \frac{1}{2}\rho(i, j) \geq \Upsilon(\rho)(i) \quad \text{a. s.}$$

For the converse inequality, we first note that

$$2v(i) \leq v(i) + v(j) + 2\varepsilon + r(i, j) = \rho(i, j) + 2\varepsilon \tag{3.2}$$

by inequality (3.1) and the definition of ρ . Moreover, for all $k \in \mathbb{N} \setminus \{i, j\}$, we obtain

$$\begin{aligned} 2v(i) - 2\varepsilon &\leq \rho(i, j) \leq \rho(i, k) \vee \rho(k, j) \\ &\leq v(k) + r(i, k) \vee r(k, j) + v(i) \vee v(j) \\ &\leq v(k) + r(i, k) + v(i) + |r(k, j) - r(i, k)| + |v(j) - v(i)| \\ &\leq \rho(i, k) + r(i, j) + 2\varepsilon \leq \rho(i, k) + 4\varepsilon \quad \text{a. s.} \end{aligned}$$

Here we use inequality (3.2) for the first and inequality (3.1) for the fifth and sixth step, the definition of ρ for the third and fifth step, and ultrametricity for the second step. By definition of the map Υ , we obtain

$$\Upsilon(\rho)(i) = \frac{1}{2} \inf_{k \in \mathbb{N} \setminus \{i\}} \rho(i, k) \geq v(i) - 3\varepsilon \quad \text{a. s.}$$

As $\varepsilon > 0$ and $i \in \mathbb{N}$ were arbitrary, it follows that $\Upsilon(\rho) = v$ a. s. \square

In Remark 3.5 below, we interpret the elements of $\hat{\mathcal{U}}$ as weighted real trees that are non-separable in general. We give a similar interpretation in Remark 3.15 in Subsection 3.4 where we have separable trees. In Remark 3.5, we use the concept of mark functions for which we refer to Depperschmidt, Greven, and Pfaffelhuber [14] and Kliem and Löhr [32]. A marked metric measure space (X, r, m) is said to admit a uniformly continuous mark function if there exists a uniformly continuous function $f : X \rightarrow \mathbb{R}_+$ such that the probability measure m on $X \times \mathbb{R}_+$ factorizes as $m(dx dv) = m_X(dx) \delta_{f(x)}(dv)$, where $m_X := m(\cdot \times \mathbb{R}_+)$. This is clearly a property of the isomorphism class.

Proposition 3.4. *Let $\chi \in \hat{\mathcal{U}}$. Then χ admits a uniformly continuous mark function.*

Proof. Let (r, v) be a random variable with the marked distance matrix distribution of χ , and let $\rho = \alpha(r, v)$. Proposition 3.3 yields $v(i) = \frac{1}{2} \inf_{k \in \mathbb{N} \setminus \{i\}} \rho(i, k)$ a. s. for all $i \in \mathbb{N}$. Hence,

$$r(i, j) = \rho(i, j) - v(i) - v(j) \geq 2 \max\{v(i), v(j)\} - v(i) - v(j) = |v(i) - v(j)| \quad \text{a. s.}$$

for all distinct $i, j \in \mathbb{N}$. The assertion follows by Lemma 1.13 of [32]. \square

Remark 3.5. Let (X, r, m) be a marked metric measure space whose isomorphism class χ lies in $\hat{\mathcal{U}}$. We assume that the closed support of the measure $m(\cdot \times \mathbb{R}_+)$ is equal to X . In this remark, we extend the space (X, r) to a real tree.

By Proposition 3.4, the marked metric measure space (X, r, m) admits a uniformly continuous mark function $f : X \rightarrow \mathbb{R}_+$. For each point $x \in X$, we take infinitely many real intervals of length $f(x)$. For each of these intervals, we glue one endpoint to x . We call the other endpoint an outer endpoint and denote it by $\ell_{x,i}$, where $i \in \mathcal{I}$ and \mathcal{I} is an index set for the intervals attached to x . The induced semi-metric ρ on the set of outer endpoints $\mathcal{L} = \{\ell_{x,i} : x \in X, i \in \mathcal{I}\}$ is given by

$$\rho(\ell_{x,i}, \ell_{y,j}) = f(x) + r(x, y) + f(y)$$

for $(x, i) \neq (y, j)$. By definition of $\hat{\mathcal{U}}$, our assumption on the support of m , and continuity of f , it follows that ρ is a semi-ultrametric on \mathcal{L} .

With (\mathcal{L}, ρ) , we associate as in Remark 1.1 a real tree on whose completion \bar{T} a probability measure is given as the image measure of $m(\cdot \times \mathbb{R}_+)$ under the isometry $X \rightarrow \bar{T}$ that is defined analogously to the map φ in Remark 2.2. This probability measure is concentrated on the subspace $\{(x, f(x)) : x \in X\}$ of the starting vertices of the external branches. Continuity of the mark function f implies that this subspace is closed.

3.3 Dust-free semi-ultrametrics and dust-free marked metric measure spaces

We call a semi-ultrametric $\rho \in \mathfrak{U}$ dust-free if $\Upsilon(\rho) = 0$. We call a marked metric measure space (X, r, m) dust-free if the probability measure is of the form $m = \mu \otimes \delta_0$ for a probability measure μ on the Borel sigma algebra on X . Then the distance matrix distribution $\alpha(\nu^{(X, r, \mu \otimes \delta_0)})$ equals the distance matrix distribution $\nu^{(X, r, \mu)}$ of the metric measure space (X, r, μ) . We call (X, r, μ) the metric measure space associated with the dust-free marked metric measure space $(X, r, \mu \otimes \delta_0)$.

Remark 3.6. Note that $f = 0$ in Proposition 3.4 if and only if χ is dust-free. Clearly, (the isomorphism class of) a marked metric measure space (X, r, m) in $\hat{\mathcal{U}}$ is dust-free if and only if a random variable with distribution $\nu^{(X, r, m)}$ is a.s. dust-free. In particular, in Theorem 1.2, the marked metric measure space χ is a.s. dust-free if and only if ρ is a.s. dust-free. In the ultrametric case, condition (4) in Vershik [47] therefore corresponds to the condition that the random semi-ultrametric is dust-free.

3.4 Marked metric measure spaces from marked distance matrices

In this subsection, we discuss the main concepts for the sampling representation in Subsection 3.5. Consider a two-step random experiment where we first sample a random (marked) metric measure space, conditionally on which we sample according to its (marked) distance matrix distribution in the second step. By Proposition 3.12 below, the realization of the (marked) metric measure space in the first step can be reconstructed from any typical realization of the (marked) distance matrix distribution in the second step. In the beginning of this subsection, we define functions by which we construct a (marked) metric measure space from a (marked) distance matrix. An interpretation of these functions is also given in Remark 3.9 below. In Remark 3.23, we state their role in the context of the ergodic decomposition.

Now we define the function $\psi : \mathfrak{D} \rightarrow \mathbb{M}$ that maps $\rho \in \mathfrak{D}$ to the isomorphism class of the metric measure space (X, ρ, μ) , given as follows: (X, ρ) is the metric completion of (\mathbb{N}, ρ) . The probability measure μ is defined as the weak limit of the probability measures $n^{-1} \sum_{i=1}^n \delta_i$ as n tends to infinity, if this weak limit exists. If the limit does not exist, we define m arbitrarily, let us set $m = \delta_1$. Furthermore, we denote by \mathfrak{D}^* the subset of distance matrices $\rho \in \mathfrak{D}$ such that the weak limit in the definition above exists.

Analogously, we define the function $\hat{\psi} : \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} \rightarrow \hat{\mathbb{M}}$ that maps (r, v) to the isomorphism class of the marked metric measure space (X, r, m) , where (X, r) is the metric completion of the semi-metric space (\mathbb{N}, r) and m is the weak limit of the probability measures

$n^{-1} \sum_{i=1}^n \delta_{(i, v_i)}$ on $X \times \mathbb{R}_+$ if this weak limit exists, else we set $m = \delta_{(1,0)}$. We denote by $\hat{\mathfrak{D}}^*$ the subset of marked distance matrices $(r, v) \in \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$ such that the weak limit in the definition above exists.

We call μ and m in the definitions of ψ and $\hat{\psi}$ also sampling measures in view of Proposition 3.10 below.

Remark 3.7. Let $(r, v) \in \mathfrak{D}$. Then $(r, v) \in \hat{\mathfrak{D}}^*$ implies $r \in \mathfrak{D}^*$. For a representative (X, r, m) of $\hat{\psi}(r, v)$, the isomorphism class of $(X, r, m(\cdot \times \mathbb{R}_+))$ equals $\psi(r)$.

Proposition 3.8. *The functions ψ and $\hat{\psi}$ are measurable.*

The proof, in which we write ψ and $\hat{\psi}$ as limits of continuous functions, is deferred to Section 11.1.

Remark 3.9 (An interpretation of ψ and $\hat{\psi}$). For $\rho \in \mathfrak{D}^* \cap \mathfrak{U}$, the probability measure in the ultrametric metric measure space $\psi(\rho)$ charges each ball with the asymptotic frequency of the corresponding block of the coalescent which is associated with ρ by (1.1).

Similarly, for $(r, v) \in \hat{\mathfrak{D}}^* \cap \hat{\mathfrak{U}}$, let (X, r, m) be the representative of $\hat{\psi}(r, v)$ from the definition of $\hat{\psi}$. We consider the completion (\bar{T}, d) of the real tree (T, d) associated with (r, v) as in Remark 2.2, and the extension $\varphi : X \rightarrow \bar{T}$ of the isometry φ from Remark 2.2. Then the image measure $\mu := \varphi(m(\cdot \times \mathbb{R}_+))$ charges each region of \bar{T} with the asymptotic frequency of the integers that label the leaves of T that are the endpoints of external branches that begin in that region.

When we sample according to the (distance) matrix distribution of the (marked) metric measure space that we have constructed from a realization of a suitable random (marked) distance matrix, we obtain a (marked) distance matrix with the same distribution, as we check in Proposition 3.10 below.

The following proposition can be compared with Lemma 8 of Vershik [47]. It will be needed also for the intertwining relations in Section 4.

Proposition 3.10. *Let (r, v) be an exchangeable random variable with values in $\hat{\mathfrak{D}}^*$. Let (r', v') be a random variable with values in $\mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$ and conditional distribution $\nu^{\hat{\psi}(r, v)}$ given $\hat{\psi}(r, v)$. Then (r', v') and (r, v) are equal in distribution.*

Remark 3.11. For an exchangeable random variable with values in \mathfrak{D}^* and a random variable ρ' with conditional distribution $\nu^{\psi(\rho)}$ given $\psi(\rho)$, the random variables ρ and ρ' are equal in distribution. This follows from Proposition 3.10, we set $(r, v) = (\rho, 0)$.

For $n \in \mathbb{N}$, we write $[n] = \{i \in \mathbb{N} : i \leq n\}$ and we denote by γ_n the restriction from $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ to $\mathbb{R}^{n^2} \times \mathbb{R}^n$, $\gamma_n(r, v) = ((r(i, j))_{i, j \in [n]}, (v(i))_{i \in [n]})$.

Proof of Proposition 3.10. Let $n \in \mathbb{N}$ and let $\phi : \mathbb{R}_+^{n^2} \times \mathbb{R}_+^n \rightarrow \mathbb{R}$ be bounded and continuous. Let (X, r, m) be the representative of $\hat{\psi}(r, v)$ as in the definition of $\hat{\mathfrak{D}}^*$. We have

$$\begin{aligned} \mathbb{E}[\phi \circ \gamma_n(r', v')] &= \mathbb{E} \left[\int m^{\otimes n}(dx \, dv'') \phi((r(x(i), x(j)))_{i, j \in [n]}, (v''(i))_{i \in [n]}) \right] \\ &= \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{\ell_1=1}^k \cdots \frac{1}{k} \sum_{\ell_n=1}^k \mathbb{E}[\phi((r(\ell_i, \ell_j))_{i, j \in [n]}, (v(\ell_i))_{i \in [n]})] \\ &= \mathbb{E}[\phi \circ \gamma_n(r, v)]. \end{aligned}$$

The second equality follows from the definition of $\hat{\mathfrak{D}}^*$ and by dominated convergence. For the third equality, we use that summands where ℓ_1, \dots, ℓ_n are not pairwise distinct vanish in the limit, and that for all other summands, the expectation in the second line equals by exchangeability the expectation in the third line. \square

Proposition 3.12. *Let $\chi \in \hat{\mathbb{M}}$ and let (r, v) be a $\mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$ -valued random variable with distribution ν^χ . Then $(r, v) \in \hat{\mathfrak{D}}^*$ a. s. and $\hat{\psi}(r, v) = \chi$ a. s.*

Remark 3.13. Proposition 3.12 is essentially Vershik's proof [47, Theorem 4] of the Gromov reconstruction theorem (where metric measure spaces are considered, cf. also [12, Theorem 1] for marked metric measure spaces). The present formulation focuses on the map $\hat{\psi}$ that will be needed in the proofs of Theorems 3.18 and 4.2 below.

Remark 3.14. For $\chi \in \mathbb{M}$ and a \mathfrak{D} -valued random variable ρ with distribution ν^χ , Proposition 3.12 implies $\rho \in \mathfrak{D}^*$ a. s. and $\psi(\rho) = \chi$ a. s.

Proof of Proposition 3.12. Let (X', r', m') be a representative of χ . W.l.o.g. we assume that the closed support of the probability measure $m'(\cdot \times \mathbb{R}_+)$ is the whole space X' , and that $(r, v) = ((r'(x(i), x(j))))_{i,j \in \mathbb{N}}, v)$ for an m' -iid sequence (x, v) . We denote by (X, r) the completion of (\mathbb{N}, r) . We endow $X' \times \mathbb{R}_+$ with the product metric $d^{X' \times \mathbb{R}_+}((x'_1, v'_1), (x'_2, v'_2)) = r'(x'_1, x'_2) \vee |v'_1 - v'_2|$, and analogously $X \times \mathbb{R}_+$. As the sequence $(x(i))_{i \in \mathbb{N}}$ is a. s. dense in X' , the isometry that maps $x(i)$ to i for all $i \in \mathbb{N}$ can a. s. be extended to a (surjective) isometry φ from X' to X . An isometry $\hat{\varphi}$ from $X' \times \mathbb{R}_+$ to $X \times \mathbb{R}_+$ is a. s. given by $(x, v') \mapsto (\varphi(x), v')$. By the Glivenko-Cantelli theorem, the probability measures $m^n := n^{-1} \sum_{i=1}^n \delta_{(x(i), v(i))}$ on $X' \times \mathbb{R}_+$ converge weakly to m' a. s. As $\hat{\varphi}$ is continuous, the probability measures $m^n := n^{-1} \sum_{i=1}^n \delta_{(i, v(i))} = \hat{\varphi}(m^n)$ on $X \times \mathbb{R}_+$ converge weakly to $m := \hat{\varphi}(m')$ a. s. This implies $(r, v) \in \mathfrak{D}^*$ a. s. and that $\hat{\psi}(r, v)$ equals the isomorphy class of (X, r, m) a. s. The second assertion follows as $\hat{\varphi}$ is a. s. a measure-preserving isometry from $X' \times \mathbb{R}_+$ to $X \times \mathbb{R}_+$, which implies that (X', r', m') and (X, r, m) have a. s. the same marked distance matrix distribution. \square

Remark 3.15 (Marked metric measure spaces and weighted real trees). Let $\chi \in \hat{\mathbb{U}}$, and let (r, v) be a $\hat{\mathbb{U}}$ -valued random variable with the marked distance matrix distribution of χ . By Proposition 3.12, we can associate a. s. with (r, v) a complete and separable weighted real tree (\bar{T}, d, μ) as in Remark 3.9. By the argument from Vershik [47, Lemma 7] which we recalled in the end of Section 1.2, the marked distance matrix (r, v) is ergodic with respect to the action of the group of finite permutations. This yields that the measure-preserving isometry class of the weighted real tree (\bar{T}, d, μ) from Remark 3.9 is constant a. s.

3.5 The sampling representation

In Proposition 3.16, we consider a typical realization of an exchangeable semi-ultrametric ρ on \mathbb{N} , and its decomposition $\beta(\rho)$. Proposition 3.16 states that the sampling measure m in the definition of $\hat{\psi}(\beta(\rho))$ in Subsection 3.4 is the weak limit of the uniform probability measures therein.

Proposition 3.16. *Let ρ be an exchangeable random variable with values in \mathfrak{U} . Then $\beta(\rho) \in \hat{\mathfrak{D}}^*$ a. s.*

In Section 11.2, we give various proofs of this proposition. In two of them, the de Finetti theorem yields the aforementioned sampling measure m as the directing measure of an exchangeable sequence.

In the dust-free case, we need not decompose the semi-metric ρ by the map β , we can work directly with the map ψ from Subsection 3.4:

Corollary 3.17. *Let ρ be an exchangeable random variable with values in \mathfrak{U} that is a.s. dust-free. Then $\rho \in \mathfrak{D}^*$ a.s.*

Proof. As ρ is a.s. dust-free, $\rho = r$ a.s. and the assertion is immediate from Proposition 3.16 and Remark 3.7. \square

We obtain a stronger version of Theorem 1.2 in which the marked metric measure space χ is stated as a deterministic function of the realization of ρ .

Theorem 3.18. *Let ρ be an exchangeable \mathfrak{U} -valued random variable. Then a $\hat{\mathfrak{U}}$ -valued random variable is defined by $\chi = \hat{\psi} \circ \beta(\rho)$. Let ρ' be a \mathfrak{U} -valued random variable with conditional distribution $\alpha(\nu^\chi)$ given χ . Then ρ and ρ' are equal in distribution and $\chi = \hat{\psi} \circ \beta(\rho')$ a.s.*

Remark 3.19. In the context of Theorem 3.18, $(\rho, \alpha(\nu^\chi))$ and $(\rho', \alpha(\nu^\chi))$ are equal in distribution. Hence, the probability kernel Λ from \mathcal{U} to \mathfrak{U} , given by $\Lambda(\nu, B) = \nu(B)$ for $\nu \in \mathcal{U}$, $B \subset \mathfrak{U}$ Borel, is a regular conditional distribution of ρ given $\alpha(\nu^\chi)$.

Proof of Theorem 3.18. The random variable χ is well-defined by Proposition 3.8. That ρ and ρ' are equal in distribution follows from Propositions 3.16 and 3.10. Proposition 3.12 implies $\chi = \hat{\psi} \circ \beta(\rho')$ a.s. \square

To deduce Theorem 1.2, it remains to show the uniqueness assertion, which follows from Proposition 3.20 below.

Proposition 3.20. *Let χ and χ' be $\hat{\mathfrak{U}}$ -valued random variables such that $\mathbb{E}[\alpha(\nu^\chi)] = \mathbb{E}[\alpha(\nu^{\chi'})]$. Then χ and χ' are equal in distribution.*

Proof. Let (r, v) be a random variable with conditional distribution ν^χ given χ , and let $\rho = \alpha(r, v)$. Propositions 3.12 and 3.3 imply $\chi = \hat{\psi} \circ \beta(\rho)$ a.s. Hence, the distribution of ρ determines the distribution of χ uniquely, which is the assertion. \square

Analogously to Theorem 3.18, we obtain using Corollary 3.17 and Remarks 3.11 and 3.14:

Corollary 3.21. *Let ρ be an exchangeable \mathfrak{U} -valued random variable that is a.s. dust-free. Then a \mathfrak{U} -valued random variable is defined by $\chi = \psi(\rho)$. Let ρ' be a \mathfrak{U} -valued random variable with conditional distribution ν^χ given χ . Then ρ and ρ' are equal in distribution and $\chi = \psi(\rho')$ a.s.*

We denote by \mathcal{U} the space of exchangeable probability distributions on \mathfrak{U} , and we endow \mathcal{U} with the Prohorov metric d_P which is complete and separable. We will also consider the subspace

$$\mathcal{U}^{\text{erg}} = \{\nu \in \mathcal{U} : \nu = \alpha(\nu^{(X, r, m)}) \text{ for some marked metric measure space } (X, r, m)\}$$

of distance matrix distributions of marked metric measure spaces. A one-to-one correspondence between the sets \mathcal{U}^{erg} and $\hat{\mathcal{U}}$ is given by the Proposition 3.3 and the definitions of these sets. Hence, also the elements of \mathcal{U}^{erg} can be seen as trees.

Remark 3.22. By Theorem 1.2, each element of \mathcal{U} is a mixture of elements of \mathcal{U}^{erg} . As the distance matrix distribution of a marked metric measure space is invariant and ergodic (with respect to the action of the group of finite permutations), and as the ergodic distributions in \mathcal{U} are extreme in the convex set \mathcal{U} (see e.g. [29, Lemma A1.2]), the set \mathcal{U}^{erg} is equal to the set of ergodic distributions in \mathcal{U} .

Remark 3.23. Also the more explicit Theorem 3.18 decomposes the distribution of the exchangeable \mathfrak{U} -valued random variable ρ' into ergodic components in the sense of e.g. Theorem A1.4 in Kallenberg [29]. The function

$$\zeta : \mathfrak{U} \rightarrow \mathcal{U}^{\text{erg}}, \quad \tilde{\rho} \mapsto \alpha(\nu^{\hat{\psi} \circ \beta(\tilde{\rho})})$$

plays the role of the decomposition map in Varadarajan [45] in the sense that $\zeta(\rho')$ is the ergodic component in whose support a typical realization ρ' lies. Note that this ergodic component is characterized by the marked metric measure space $\chi = \hat{\psi} \circ \beta(\rho')$, and in the dust-free case also by the metric measure space $\psi(\rho')$. Some further references on the ergodic decomposition are given e.g. in [29, p.475].

The following corollary, which will be applied in Section 12, corresponds to the uniqueness of the ergodic component of a typical realization of ρ' in the context of Remark 3.23.

Corollary 3.24. *Let $\nu \in \mathcal{U}^{\text{erg}}$ and let ρ be a random variable with distribution ν . Then the distance matrix distribution of $\hat{\psi} \circ \beta(\rho)$ equals ν a. s.*

Proof. By definition of \mathcal{U}^{erg} , there exists $\chi \in \hat{\mathcal{U}}$ with distance matrix distribution $\nu^\chi = \nu$. Propositions 3.3 and 3.12 imply $\hat{\psi} \circ \beta(\rho) = \chi$ a. s. \square

The following corollary to Theorem 1.2 implies that $(\mathcal{U}^{\text{erg}}, d_P)$ is Polish which will be applied in [26].

Corollary 3.25. *The subspace \mathcal{U}^{erg} is closed in \mathcal{U} .*

Proof. Let $(\nu^n, n \in \mathbb{N}) \subset \mathcal{U}^{\text{erg}}$ be a sequence that converges to some ν in (\mathcal{U}, d_P) . By Theorem 1.2, there exists a random marked metric measure space χ with $\nu = \mathbb{E}[\alpha(\nu^\chi)]$. We show that $\nu' := \alpha(\nu^\chi)$ is independent of itself. With this property, it follows that $\nu' = \nu$ a. s. and $\nu \in \mathcal{U}^{\text{erg}}$.

To show this independence, let $\Psi_1, \Psi_2 : \mathcal{U} \rightarrow \mathbb{R}$ be bounded and continuous. As also the function $\Psi_1 \Psi_2$ is bounded and continuous,

$$\mathbb{E}[\Psi_1(\nu') \Psi_2(\nu')] = \lim_{n \rightarrow \infty} \mathbb{E}[\Psi_1(\nu^n) \Psi_2(\nu^n)] = \mathbb{E}[\Psi_1(\nu')] \mathbb{E}[\Psi_2(\nu')].$$

We conclude by [18, Theorem 3.4.6]. \square

4 Application to tree-valued processes

Using the function $\hat{\psi}$ from Section 3.4, we map a Markov process whose states are exchangeable \mathfrak{U} -valued random variables to a process with values in the space of isomorphism classes of marked metric measure spaces. At each time, the state of the image process is the marked metric measure space from the representation (Theorem 3.18) of the state of the \mathfrak{U} -valued process. We also consider the process of the distance matrix distributions of these marked metric measure spaces. In the dust-free case, we can also work with isomorphism classes of metric measure spaces and the map ψ as in Corollary 3.21.

Using the criterion of Rogers and Pitman [42, Theorem 2], we show that also the image process is Markovian. A martingale problem for the \mathfrak{U} -valued process yields a martingale problem for the image process.

Remark 4.1. Theorem 4.2 below is an example for Markov mapping. That the image processes solve the martingale problems given in Theorem 4.2 can also be seen as a consequence of Lemma A.2 in Kurtz and Nappo [34]. Uniqueness of these martingale problems follows under suitable conditions from Lemma 3.5 in Kurtz [33]. In Section 8, we show uniqueness directly by duality for concrete examples.

The so-called polynomials and marked polynomials, introduced in [12, 24] have been used as domains of martingale problems in e. g. [13, 14, 25]. We recall them here, adapting the definition to our present use of the marks. The uniform continuity of the derivative in the definitions of \mathcal{C}_n and $\hat{\mathcal{C}}_n$ below will turn out useful in [26]. Recall the restriction γ_n from $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ to $\mathbb{R}^{n^2} \times \mathbb{R}^n$ and the notation $[n] = \{1, \dots, n\}$ for $n \in \mathbb{N}$. We denote also by γ_n the restriction from $\mathbb{R}^{\mathbb{N}^2}$ to \mathbb{R}^{n^2} , $\gamma_n(r, v) = ((r_{i,j})_{i,j \in [n]}, (v_i)_{i \in [n]})$. Let \mathcal{C}_n denote the set of bounded differentiable functions $\mathbb{R}^{n^2} \rightarrow \mathbb{R}$ with bounded uniformly continuous derivative. For $\phi \in \mathcal{C}_n$, we denote also by ϕ the function $\phi \circ \gamma_n : \mathbb{R}^{\mathbb{N}^2} \rightarrow \mathbb{R}$, and we call the function $\mathbb{U} \rightarrow \mathbb{R}$, $\chi \mapsto \nu^\chi \phi$ the polynomial associated with ϕ . Similarly, we denote by $\hat{\mathcal{C}}_n$ the set of bounded differentiable functions $\mathbb{R}^{n^2} \times \mathbb{R}^n \rightarrow \mathbb{R}$ with uniformly continuous derivative. For $\phi \in \hat{\mathcal{C}}_n$, we denote also by ϕ the function $\phi \circ \gamma_n : \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$, and we call the function $\hat{\mathbb{U}} \rightarrow \mathbb{R}$, $\chi \mapsto \nu^\chi \phi$ the marked polynomial associated with ϕ . We write $\mathcal{C} = \bigcup_{n \in \mathbb{N}} \mathcal{C}_n$ and $\hat{\mathcal{C}} = \bigcup_n \hat{\mathcal{C}}_n$. We denote the set of polynomials by

$$\Pi = \{\mathbb{U} \rightarrow \mathbb{R}, \chi \mapsto \nu^\chi \phi : \phi \in \mathcal{C}\},$$

the set of marked polynomials by

$$\hat{\Pi} = \{\hat{\mathbb{U}} \rightarrow \mathbb{R}, \chi \mapsto \nu^\chi \phi : \phi \in \hat{\mathcal{C}}\},$$

and we define the set of test functions

$$\mathcal{C} = \{\mathcal{U}^{\text{erg}} \rightarrow \mathbb{R}, \nu \mapsto \nu \phi : \phi \in \mathcal{C}\}.$$

For a metric space E , a subset \mathcal{D} of the set $M_b(E)$ of bounded measurable functions $E \rightarrow \mathbb{R}$, and an operator $G : \mathcal{D} \rightarrow M_b(E)$, we mean by a solution of the martingale problem (G, \mathcal{D}) a progressive E -valued process $(X_t, t \in \mathbb{R}_+)$ such that for every $f \in \mathcal{D}$, the process

$$f(X_t) - \int_0^t Gf(X_s) ds$$

is a martingale with respect to the filtration induced by $(X_t, t \in \mathbb{R}_+)$, cf. Ethier and Kurtz [18, p. 173].

Theorem 4.2. *Let $(\rho_t, t \in \mathbb{R}_+)$ be a \mathfrak{U} -valued Markov process. Assume that for each $t \in \mathbb{R}_+$, the random variable ρ_t is exchangeable. Let $A : \mathcal{C} \rightarrow M_b(\mathbb{R}^{\mathbb{N}^2})$ and $\hat{A} : \hat{\mathcal{C}} \rightarrow M_b(\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}})$ be operators. Define the \mathbb{U} -valued process $(\chi_t, t \in \mathbb{R}_+) := (\psi(\rho_t), t \in \mathbb{R}_+)$, the $\hat{\mathbb{U}}$ -valued process $(\hat{\chi}_t, t \in \mathbb{R}_+) := (\hat{\psi}(\beta(\rho_t)), t \in \mathbb{R}_+)$, and the \mathcal{U}^{erg} -valued process $(\nu_t, t \in \mathbb{R}_+) := (\alpha(\nu^{\hat{\chi}_t}), t \in \mathbb{R}_+)$. Then the following two assertions hold:*

- (i) *The process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ is Markovian. If the $\hat{\mathfrak{U}}$ -valued process $(\beta(\rho_t), t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{A}, \hat{\mathcal{C}})$, then $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{B}, \hat{\Pi})$, given by*

$$\hat{B}\Phi(\chi) = \nu^\chi(\hat{A}\phi)$$

for all $\phi \in \hat{\mathcal{C}}$ with associated polynomial Φ , and all $\chi \in \hat{\mathbb{U}}$.

- (ii) *The process $(\nu_t, t \in \mathbb{R}_+)$ is Markovian. If $(\rho_t, t \in \mathbb{R}_+)$ solves the martingale problem (A, \mathcal{C}) , then $(\nu_t, t \in \mathbb{R}_+)$ solves the martingale problem (C, \mathcal{C}) , given by*

$$C\Psi(\nu) = \nu(A\phi)$$

for all $\nu \in \mathcal{U}^{\text{erg}}$ and $\phi \in \mathcal{C}$, and the function $\Psi \in \mathcal{C}$, $\nu' \mapsto \nu'\Psi$.

Assertion (iii) below holds under the additional assumption that ρ_t is a.s. dust-free for each $t \in \mathbb{R}_+$.

- (iii) *The process $(\chi_t, t \in \mathbb{R}_+)$ is Markovian. If $(\rho_t, t \in \mathbb{R}_+)$ solves the martingale problem (A, \mathcal{C}) , then $(\chi_t, t \in \mathbb{R}_+)$ solves the martingale problem (B, Π) , given by*

$$B\Phi(\chi) = \nu^\chi(A\phi)$$

for all $\phi \in \mathcal{C}$ with associated polynomial Φ , and all $\chi \in \mathbb{U}$.

Remark 4.3. The process $(\beta(\rho_t), t \in \mathbb{R}_+)$ in Theorem 4.2 is also Markov. This follows as $\rho_t = \alpha(\beta(\rho_t))$ and as $(\rho_t, t \in \mathbb{R}_+)$ is Markov by assumption.

Remark 4.4. In Theorem 4.2, if ρ_t is dust-free for some $t \in \mathbb{R}_+$, then χ_t is the (isomorphism class of the) metric measure space associated (as in Section 3.3) with (any representative of) the dust-free marked metric measure space $\hat{\chi}_t$, and we have $\nu_t = \nu^{\chi_t}$. The process $(\chi_t, t \in \mathbb{R}_+)$ is only relevant in the dust-free case: If ρ_t is not dust-free, then $\psi(\rho_t)$ is a.s. just the arbitrary element of \mathbb{M} from the definition of ψ in Section 3.4.

Remark 4.5. In particular in Sections 9 – 10 and in [26], we need convergence determining sets of test functions. As in [12, 24, 36], the sets Π and $\hat{\Pi}$ are convergence determining in \mathbb{U} and $\hat{\mathbb{U}}$, respectively. The argument from [36, Corollary 2.8] also applies for \mathcal{C} : The algebra \mathcal{C} generates the product topology on $\mathbb{R}^{\mathbb{N}^2}$. By a theorem due to Le Cam, see e.g. [36, Theorem 2.7] and the references therein, it follows that \mathcal{C} is convergence determining in \mathfrak{U} . Hence, \mathcal{C} generates the weak topology on \mathcal{U}^{erg} . As $\hat{\Pi}$ is an algebra (see [12, 24]) and by definition of \mathcal{U}^{erg} , also \mathcal{C} is an algebra. Again by [36, Theorem 2.7], it follows that \mathcal{C} is convergence determining in \mathcal{U}^{erg} .

Remark 4.6. The set of polynomials $\Pi' = \{\hat{\mathbb{U}} \rightarrow \mathbb{R}, \chi \mapsto \alpha(\nu^\chi)\phi : \phi \in \mathcal{C}\}$ is separating on $\hat{\mathbb{U}}$. This follows from Propositions 3.3 and 3.12 as in the proof of Proposition 3.20. Nevertheless, we work with the space $\hat{\Pi}$ of test functions on $\hat{\mathbb{M}}$ as Π' is not convergence determining, a counterexample can be constructed from [24, Example 2.12(ii)].

The following property is central in the proof of Theorem 4.2.

Proposition 4.7. *Let $t \in \mathbb{R}_+$, and let $f : \mathfrak{U} \rightarrow \mathbb{R}$, $g : \hat{\mathfrak{U}} \rightarrow \mathbb{R}$ be bounded measurable functions. If the assumptions of Theorem 4.2 hold, then*

$$\mathbb{E}[g(\beta(\rho_t))] = \mathbb{E}[\nu^{\hat{\chi}_t} g]$$

and

$$\mathbb{E}[f(\rho_t)] = \mathbb{E}[\nu_t f].$$

If the assumptions of Theorem 4.2 hold and ρ_t is a.s. dust-free, then

$$\mathbb{E}[f(\rho_t)] = \mathbb{E}[\nu^{\chi_t} f].$$

Proof. This is immediate from Propositions 3.16 and 3.10, the definition of ν_t , Corollary 3.17, and Remark 3.11. \square

Remark 4.8. In the context of Theorem 4.2(i), let $(P_t, t \in \mathbb{R}_+)$ denote the semigroup on $M_b(\hat{\mathfrak{U}})$ of the Markov process $(\beta(\rho_t), t \in \mathbb{R}_+)$, and let $(Q_t, t \in \mathbb{R}_+)$ denote the semigroup on $M_b(\hat{\mathfrak{U}})$ of the Markov process $(\hat{\chi}_t, t \in \mathbb{R}_+)$. Let K denote the probability kernel from $\hat{\mathfrak{U}}$ to $\hat{\mathfrak{U}}$, given by $K(\chi, \cdot) = \nu^\chi$ for $\chi \in \hat{\mathfrak{U}}$. Then Proposition 4.7 yields the intertwining relation $Q_t K = K P_t$ which is condition (b) in [42, Theorem 2]. Many papers appeared on intertwining of Markov processes, a classical one is for instance [9].

Proof of Theorem 4.2. We apply [42, Theorem 2] to the semigroup of the Markov process $(\beta(\rho_t), t \in \mathbb{R}_+)$, the measurable map $\hat{\psi} : \hat{\mathfrak{U}} \rightarrow \hat{\mathfrak{U}}$, and the kernel K from $\hat{\mathfrak{U}}$ to $\hat{\mathfrak{U}}$ given by $K(\chi, \cdot) = \nu^\chi$. Clearly, Theorem 2 in [42] also holds when the initial state y therein is random. Then by Proposition 4.7, condition (b) and the condition on the initial state in [42, Theorem 2] are satisfied. Condition (a) in [42, Theorem 2] follows from Proposition 3.12 as $f(\chi) = \nu^\chi(f \circ \hat{\psi})$ for all $\chi \in \hat{\mathfrak{U}}$ and all bounded measurable $f : \hat{\mathfrak{U}} \rightarrow \mathbb{R}$. The Markov property of $(\hat{\chi}_t, t \in \mathbb{R}_+)$ now follows from [42, Theorem 2].

Now we prove that $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{B}, \hat{\Pi})$. If $(\beta(\rho_t), t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{A}, \hat{\mathcal{C}})$ in (i), then for all $\phi \in \hat{\mathcal{C}}$ with associated marked polynomial Φ ,

$$\begin{aligned} 0 &= \mathbb{E}[\phi(\beta(\rho_t)) - \phi(\beta(\rho_0)) - \int_0^t \hat{A}\phi(\beta(\rho_u)) du] \\ &= \mathbb{E}[\nu^{\hat{\chi}_t} \phi] - \mathbb{E}[\nu^{\hat{\chi}_0} \phi] - \int_0^t \mathbb{E}[\nu^{\hat{\chi}_u}(\hat{A}\phi)] du \\ &= \mathbb{E}[\Phi(\hat{\chi}_t) - \Phi(\hat{\chi}_0) - \int_0^t \hat{B}\Phi(\hat{\chi}_u) du] \end{aligned} \tag{4.1}$$

by Proposition 4.7, Fubini, and the definitions Φ and \hat{B} . By the Markov property of $(\chi_s, s \in \mathbb{R}_+)$ and equation (4.1), it now follows for all $s \in \mathbb{R}_+$ and all $(\hat{\chi}_u, u \in [0, s])$ -measurable events A that

$$\mathbb{E}[\Phi(\hat{\chi}_{s+t}) - \Phi(\hat{\chi}_s) - \int_s^{s+t} B\Phi(\hat{\chi}_u) du; A] = 0$$

which shows assertion (i).

The proof of (ii) is analogous, we apply [42, Theorem 2] to the Markov process $(\rho_t, t \in \mathbb{R}_+)$, the measurable map $\mathfrak{U} \rightarrow \mathcal{U}^{\text{erg}}$, $\rho \mapsto \alpha(\nu^{\hat{\psi}(\beta(\rho))})$, and the probability kernel from \mathcal{U}^{erg} to \mathfrak{U} given by $(\nu, B) \mapsto \nu(B)$. In particular, condition (a) in [42, Theorem 2] is satisfied by Propositions 3.12 and 3.3, and by definition of \mathcal{U}^{erg} .

Also the proof of (iii) is analogous. We apply [42, Theorem 2] to the process $(\rho_t, t \in \mathbb{R}_+)$, the measurable map $\psi : \mathfrak{U} \rightarrow \mathbb{U}$, and the probability kernel from \mathbb{U} to \mathfrak{U} given by $(\chi, B) \mapsto \nu^\chi(B)$. We use the assumption that ρ_t is a.s. dust-free in the application of Proposition 4.7 and Remark 3.14. \square

5 Genealogy in the lookdown model

In this and the next section, we define a Markov process $(\rho_t, t \in \mathbb{R}_+)$ to which we will later apply Theorem 4.2. We read off this process from a population model that is driven by a deterministic point measure in this section and by a Poisson random measure in the next section. We remark that for the lookdown model of Donnelly and Kurtz [15], the process of the evolving genealogical distances and its martingale problem are considered in Remark 2.20 of Greven, Pfaffelhuber, and Winter [25].

We denote by \mathcal{P} the set of partitions of \mathbb{N} . We endow \mathcal{P} with the topology in which a sequence of partitions converges if and only if the sequences of their finite restrictions converge. For $n \in \mathbb{N}$, we denote by \mathcal{P}_n the set of partitions of $[n] = \{1, \dots, n\}$. We denote the restriction map from \mathcal{P} to \mathcal{P}_n by γ_n , that is, $\gamma_n(\pi) = \{B \cap [n] : B \in \pi\} \setminus \{\emptyset\}$. Recall that other restriction maps, e.g. from $\mathbb{R}^{\mathbb{N}^2} \rightarrow \mathbb{R}^{n^2}$ are also denoted by γ_n . Moreover, we denote by $\mathbf{0}_n = \{\{1\}, \dots, \{n\}\}$ the partition in \mathcal{P}_n that consists of singletons only, and by $\mathcal{P}^n = \{\pi \in \mathcal{P} : \gamma_n(\pi) \neq \mathbf{0}_n\}$ the set of partitions of \mathbb{N} in which the first n integers are not all in different blocks. Furthermore, for $\pi \in \mathcal{P}$, we denote by $B_1(\pi), B_2(\pi), \dots$ the enumeration of the blocks of π with $\min B_1(\pi) < \min B_2(\pi) < \dots$. For $i \in \mathbb{N}$, we denote by $\pi(i)$ the integer j that satisfies $i \in B_j(\pi)$.

We use a lookdown model as the population model. In this model, there are countably infinitely many levels which are labeled by \mathbb{N} , and each level is occupied by one particle at each time $t \in \mathbb{R}_+$. The particles undergo reproduction events which are encoded by a simple point measure η on $(0, \infty) \times \mathcal{P}$. A simple point measure is a purely atomic measure whose atoms all have mass 1. Let us impose a further assumption on η , namely

$$\eta((0, t] \times \mathcal{P}^n) < \infty \quad \text{for all } t \in (0, \infty) \text{ and } n \in \mathbb{N}. \quad (5.1)$$

The interpretation of a point (t, π) of η is that the following reproduction event occurs: At time $t-$, the particles on the levels $i \in \mathbb{N}$ with $i > \#\pi$ are removed. At time t , for each $i \in [\#\pi]$, the particle that was on level i at time $t-$ assumes level $\min B_i(\pi)$ and has offspring on all other levels in $B_i(\pi)$. Thus, the level of a particle is non-decreasing as time evolves. Condition (5.1) means that for each $n \in \mathbb{N}$, only finitely many particles jump away from the first n levels in bounded time intervals.

For all $0 \leq s \leq t$, each particle at time t has an ancestor at time s . We denote by $A_s(t, i)$ the level of the ancestor at time s of the particle on level i at time t such that the maps $s \mapsto A_s(t, i)$ and $t \mapsto A_s(t, i)$ are càdlàg. Then $A_s(t, i)$ is well-defined as $s \mapsto A_{t-s}(t, i)$ is non-increasing.

Remark 5.1. We will use that the trajectories of the particles are non-crossing in the following sense: For any times $s \leq t$ and particles x, y on levels $i_x \leq i_y$ at time $s \in \mathbb{R}_+$, particle x is still alive if particle y is still alive, in which case the particles x and y occupy levels $j_x \leq j_y$. In particular, if infinitely many particles at time s survive until time t , then all particles at time s survive until time t .

We are interested in the process of the genealogical distances between the particles that live at the respective times. Let $\rho_0 \in \mathbb{R}^{\mathbb{N}^2}$. (We can assume $\rho_0 \in \mathfrak{U}$ here, but differentiability will be more elementary in the larger space, as a matter of taste.) We interpret $\rho_0(i, j)$ as the genealogical distance between the particles on levels i and j at time 0. We define the genealogical distance between the particles on levels i and j at time t by

$$\rho_t(i, j) = \begin{cases} 2t - 2 \sup\{s \in [0, t] : A_s(t, i) = A_s(t, j)\} & \text{if } A_0(t, i) = A_0(t, j) \\ 2t + \rho_0(A_0(t, i) + A_0(t, j)) & \text{else.} \end{cases}$$

In words, the genealogical distance between two particles at a fixed time is twice the time back to their most recent common ancestor, if such an ancestor exists, else it is given by the genealogical distance between the ancestors at time zero.

Remark 5.2. If $\rho_0 \in \mathfrak{U}$, then $\rho_t \in \mathfrak{U}$ for each $t \in \mathbb{R}_+$. Indeed, a semi-metric ρ on \mathbb{N} is a semi-ultrametric if and only if for each $s \in \mathbb{R}_+$, an equivalence relation \sim on \mathbb{N} is given by $i \sim j :\Leftrightarrow \rho(i, j) \leq s$. If this property holds for ρ_0 , then the definition of ρ_t readily yields that it also holds for ρ_t .

In the remainder of this section, we describe the process $(\rho_t, t \in \mathbb{R}_+)$ in a more formal way which will be useful for the description by martingale problems in Section 6. With each partition $\pi \in \mathcal{P}_n$ we associate a transformation $\mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$, which we also denote by π , by

$$\pi(\rho) = (\rho(\pi(i), \pi(j)))_{i, j \in [n]}. \quad (5.2)$$

Here $\pi(i)$ denotes the integer k such that i is in the k -th block, when blocks are ordered according to their minimal elements. Note that for each reproduction event encoded by a point $(t, \pi) \in \eta$, the jump of the process $(\rho_t, t \in \mathbb{R}_+)$ can be described by

$$\gamma_n(\pi)(\gamma_n(\rho_{t-})) = \gamma_n(\rho_t). \quad (5.3)$$

In particular, $\gamma_n(\pi) = \mathbf{0}_n$ if $\pi \in \mathcal{P} \setminus \mathcal{P}^n$, and $\mathbf{0}_n$ acts as the identity on \mathbb{R}^{n^2} . By assumption (5.1), there are only finitely many reproduction events in bounded time intervals that result in a jump of the process $(\gamma_n(\rho_t(i, j)), t \in \mathbb{R}_+)$. Between such jumps, the genealogical distances grow linearly with slope 2, that is, $\rho_t(i, j) + 2s = \rho_{t+s}(i, j)$ for distinct $i, j \in [n]$ and $t, s \in \mathbb{R}_+$ with $\eta((t, t+s] \times \mathcal{P}^n) = 0$.

Remark 5.3. Schweinsberg [44] constructs the Ξ -coalescent analogously from a point measure. The population model described in this section can be seen as the population model that underlies the dual flow of partitions in Foucart [23]. A lookdown model with a reproduction mechanism that is different in the case with simultaneous multiple reproduction events is studied by Birkner et al. [8]. In this model, a partition $\pi \in \mathcal{P}$ encodes the following reproduction event: Let $i_1 < i_2 < \dots$ be the increasing enumeration of the integers that either form singletons or are non-minimal elements of blocks of π . For each

$j \in \mathbb{N}$, the particle on level i_j moves to the level given by the j -th lowest singleton of π if π has at least j singletons, else the particle is removed. For each non-singleton block $B \in \pi$, the particle on level $\min B$ remains on its level and has one offspring on each level in $B \setminus \{\min B\}$. Here the trajectories of the particles may cross: Consider a partition $\pi \in \mathcal{P}$ such that 1 and 2 are in the same block, 4 forms a singleton, and 3 is the minimal element of a non-singleton block. If the reproduction event encoded by π occurs at time $t \in (0, \infty)$, then there exists $s \in (0, t)$ such that the particle on level 3 at time s is on level 3 also at time t , and the particle on level 2 at time s jumps to level 4 at time t . Such a crossing cannot occur in our population model by Remark 5.1.

6 The Ξ -lookdown model

The population model from Section 5 will be driven by a Poisson random measure η on $(0, \infty) \times \mathcal{P}$ as in Schweinsberg [44], Bertoin [4], and Foucart [23].

To define this Poisson random measure, we briefly recall Kingman's correspondence. For a full account, see e.g. [4, Section 2.3.2]. Kingman's correspondence is a one-to-one correspondence between the distributions of the exchangeable random partitions of \mathbb{N} and the probability measures on the simplex

$$\Delta = \{x = (x_1, x_2, \dots) : x_1 \geq x_2 \geq \dots \geq 0, |x|_1 \leq 1\},$$

where $|x|_p = (\sum_{i \in \mathbb{N}} x_i^p)^{1/p}$. Every $x \in \Delta$ can be interpreted as a partition of $[0, 1]$ into subintervals of lengths x_1, x_2, \dots , and possibly another interval of length $1 - |x|_1$ which may be called the dust interval. Let U_1, U_2, \dots be iid uniform random variables with values in $[0, 1]$. The paintbox partition associated with x is the exchangeable random partition of \mathbb{N} where two different integers i and j are in the same block if and only if U_i and U_j fall into a common subinterval that is not the dust interval. This construction defines a probability kernel κ from Δ to \mathcal{P} . Conversely, every exchangeable random partition π in \mathcal{P} has distribution $\int_{\Delta} \nu(dx) \kappa(x, \cdot)$ for some distribution ν on Δ . Here x is the random vector in Δ of the asymptotic frequencies of the blocks of π .

Let Ξ be a finite measure on Δ . We decompose

$$\Xi = \Xi_0 + \Xi\{0\}\delta_0. \tag{6.1}$$

For $i, j \in \mathbb{N}$ with $i \neq j$, we denote by $K_{i,j}$ the partition in \mathcal{P} that contains the block $\{i, j\}$ and apart from that only singleton blocks. We define a σ -finite measure H_{Ξ} on \mathcal{P} by

$$H_{\Xi}(d\pi) = \int_{\Delta} \kappa(x, d\pi) |x|_2^{-2} \Xi_0(dx) + \Xi\{0\} \sum_{1 \leq i < j} \delta_{K_{i,j}}(d\pi).$$

Let η be a Poisson random measure on $(0, \infty) \times \mathcal{P}$ with intensity $dt H_{\Xi}(d\pi)$. In Section 7.2, we will see that η satisfies a.s. condition (5.1), see also Remark 6.1 below. Hence, we can define the population model from Section 5 from almost every realization of η and every $\rho_0 \in \mathbb{R}^{\mathbb{N}^2}$.

From the description around equation (5.3) in Section 5 and the properties of Poisson random measures, it follows that each of the processes $(\gamma_n(\rho_t), t \in \mathbb{R}_+)$, $n \in \mathbb{N}$, is Markov, hence also the process $(\rho_t, t \in \mathbb{R}_+)$. For each $n \in \mathbb{N}$ and $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$, the rate at which

reproduction events encoded by partitions in $\gamma_n^{-1}(\pi) = \{\pi' \in \mathcal{P} : \gamma_n(\pi') = \pi\}$ occur in the lookdown model is given by $\lambda_\pi = H_\Xi(\gamma_n^{-1}(\pi))$. The rates λ_π are calculated explicitly in (7.4) and (7.3) in Section 7.2.

Remark 6.1. The quantity λ_π is the coagulation rate q_π in Section 4.2.1 of Bertoin [4]. It is related to the quantity $\lambda_{n;k_1,\dots,k_r;s}$ from Schweinsberg [44] by $\lambda_\pi = \lambda_{n;k_1,\dots,k_r;s}$, where k_1, \dots, k_r denote the sizes of the non-singleton blocks of π , and $s = n - k_1 - \dots - k_r$. This can be seen by a comparison of equations (7.4) and (7.3) with equation (11) in [44]. In particular, equation (18) in [44] implies that η satisfies a.s. condition (5.1).

In the next proposition, we state a martingale problem for the process $(\rho_t, t \in \mathbb{R}_+)$ from Section 5, driven by the Poisson random measure η .

Recall the set \mathcal{C} from Section 4. For $\phi \in \mathcal{C}$ and $\rho \in \mathbb{R}^{\mathbb{N}^2}$, we write

$$\langle \nabla \phi, \underline{\underline{2}} \rangle(\rho) = 2 \sum_{\substack{i,j \in \mathbb{N} \\ i \neq j}} \frac{\partial}{\partial \rho(i,j)} \phi(\rho).$$

Proposition 6.2. *Define an operator $A = A_{\text{grow}} + A_{\text{repr}}$ with domain \mathcal{C} by*

$$A_{\text{grow}}\phi(\rho) = \langle \nabla \phi, \underline{\underline{2}} \rangle(\rho)$$

and

$$A_{\text{repr}}\phi(\rho) = \sum_{\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}} \lambda_\pi(\phi(\pi(\gamma_n(\rho))) - \phi(\rho))$$

for $n \in \mathbb{N}$, $\phi \in \mathcal{C}_n$, and $\rho \in \mathbb{R}^{\mathbb{N}^2}$. Then the stochastic process $(\rho_t, t \in \mathbb{R}_+)$ solves the martingale problem (A, \mathcal{C}) .

Proposition 6.2 follows from the discussion above and the description of the process $(\gamma_n(\rho_t), t \in \mathbb{R}_+)$ in Section 5. As in [25], the operator A_{grow} reflects the growth of the genealogical distances between reproduction events that affects them. The operator A_{repr} stands for the jumps of the genealogical distances in reproduction events, as described by equation (5.3).

6.1 Properties of the genealogy at a fixed time

We assume that the process $(\rho_t, t \in \mathbb{R}_+)$ is obtained from the Poisson random measure η and a $\mathbb{R}^{\mathbb{N}^2}$ -valued random variable ρ_0 . To apply Theorem 4.2, we need exchangeability of the random variable ρ_t for each $t \in \mathbb{R}_+$.

Proposition 6.3. *Let $t \in \mathbb{R}_+$ and assume that ρ_0 is exchangeable. Then ρ_t is exchangeable.*

Remark 6.4. For $t \in \mathbb{R}_+$, let $(\Pi_s^{(t)}, s \in [0, t])$ be the \mathcal{P} -valued stochastic process such that two integers $i, j \in \mathbb{N}$ are in the same block of $\Pi_s^{(t)}$ if and only if $\rho_t(i, j) \leq 2s$. Then a comparison of the Poisson process construction of the Ξ -coalescent in [44, Section 3] with the Poisson process construction from Sections 5 and 6 shows that a Ξ -coalescent up to time t is given by the process $(\Pi_s^{(t)}, s \in [0, t])$. The distance matrix $\rho_t \wedge (2t)$ can be retrieved from $(\Pi_s^{(t)}, s \in [0, t])$ by

$$\rho_t(i, j) \wedge (2t) = 2 \inf\{s \in [0, t] : i \text{ and } j \text{ are in the same block of } \Pi_s^{(t)}, \text{ or } s = t\}$$

As Ξ -coalescents are exchangeable, it follows that the random variable $(\rho_t(i, j) \wedge (2t))_{i, j \in \mathbb{N}}$ is exchangeable. We remark that the collection of partitions $(\Pi_{(t-s)-}^{(t)}, 0 \leq s \leq t)$ is the dual flow of partitions from Foucart [23] in one-sided time. We also remark that preservation of exchangeability in the lookdown model is studied in e. g. [8, 15, 16].

To prove Proposition 6.3, we show in Lemma 6.5 below that exchangeability is preserved in single reproduction events. Then we construct the state ρ_t , restricted to the first $n \in \mathbb{N}$ particles, from the state ρ_0 and the reproduction events before time t that affect the genealogical distances between the first n individuals. Here we use the description from the end of Section 5.

For $n \in \mathbb{N}$, we define the action of the group S_n of permutations of $[n]$ on \mathcal{P}_n and \mathbb{R}^{n^2} , respectively, by

$$p(\pi) = \{p(B) : B \in \pi\} \quad \text{and} \quad p(\rho) = (\rho(p(i), p(j)))_{i, j \in [n]}$$

for each $p \in S_n$, $\pi \in \mathcal{P}_n$, $\rho \in \mathbb{R}^{n^2}$. A random variable with values for instance in the space \mathcal{P}_n of partitions of $[n]$ or in \mathbb{R}^{n^2} is called exchangeable if its distribution is invariant under the action of S_n .

Lemma 6.5. *Let $n \in \mathbb{N}$, let π be an exchangeable random partition of $[n]$, and let ρ be an exchangeable random variable with values in \mathbb{R}^{n^2} . Assume that π and ρ are independent. Then the random variable $\pi(\rho)$ is exchangeable.*

Lemma 6.5 can be seen as a generalization of Lemma 4.3 of Bertoin [4].

Proof. Let $p \in S_n$. For each partition $\pi' \in \mathcal{P}_n$, the blocks of π' are in one-to-one correspondence with the blocks of $p(\pi')$ via the bijection that maps a block $B \in \pi'$ to the block $p(B) \in p(\pi')$. Also, the blocks of π' are in one-to-one correspondence with the integers in $[n]$ that are the minimal elements of the blocks of π' . The same holds for the blocks of $p(\pi')$ and their minimal elements. It follows that the minimal elements of the blocks of π' are in one-to-one correspondence with the minimal elements of the blocks of $p(\pi')$. We extend this one-to-one correspondence arbitrarily to a bijection from $[n]$ to itself which we denote by $f(\pi')$. This defines a map $f : \mathcal{P}_n \rightarrow S_n$ which satisfies

$$\pi'(i) = f(\pi')(p(\pi')(p(i))) \tag{6.2}$$

for all $\pi' \in \mathcal{P}_n$ and $i \in [n]$. This equation holds as $\pi'(i)$, by its definition in Section 5, is a minimal element of a block of π' and as $p(\pi')(p(i))$ is the minimal element of the corresponding block of π' . By the definition (5.2) of the transformation on \mathbb{R}^{n^2} associated with each element of \mathcal{P}_n , equation (6.2) implies

$$\pi'(\rho') = p(p(\pi')(f(\pi')(\rho'))) \tag{6.3}$$

for all $\pi' \in \mathcal{P}_n$ and $\rho' \in \mathbb{R}^{n^2}$.

By assumption, $f(\pi)$ and π are equal in distribution. As π and ρ are independent, $f(\pi)(\rho)$ and ρ are equal in distribution. As the distribution of $f(\pi')(\rho)$ is the same for all $\pi' \in \mathcal{P}_n$, namely equal to the distribution of ρ , it follows that $f(\pi)(\rho)$ and π are independent. This implies that $\pi(\rho)$ and $p(\pi)(f(\pi)(\rho))$ are equal in distribution. The assertion follows from equation (6.3). \square

Proof of Proposition 6.3. Let $n \in \mathbb{N}$. For $s \in \mathbb{R}_+$, we define the map

$$\lambda_s : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}, \quad \rho' \mapsto \rho_n + \underline{\underline{2}}_n s,$$

where $\underline{\underline{2}}_n = 2(\mathbf{1}\{i \neq j\})_{i,j \in [n]}$. We will use the map λ_s to account for the linear growth of the genealogical distances between reproduction events.

On an event of probability 1, let $(t_1, \pi_1), (t_2, \pi_2), \dots$ be the points of η in $(0, t] \times \mathcal{P}^n$. Let $L = \eta((0, t] \times \mathcal{P}_n)$. Conditionally given (t_1, \dots, t_L) , the partitions π_1, \dots, π_L are independent and for each $k \in \mathbb{N}$, the restriction $\gamma_n(\pi_k)$ is exchangeable. This follows from the properties of Poisson random measures and the definition of η . From the description in Section 5, we have

$$\gamma_n(\rho_t) = \lambda_{t-t_L} \circ \gamma_n(\pi_L) \circ \lambda_{t_L-t_{L-1}} \circ \dots \circ \gamma_n(\pi_1) \circ \lambda_{t_1}(\gamma_n(\rho_0)) \quad \text{a.s.}$$

on the event $\{L \geq 1\}$, and $\gamma_n(\rho_t) = \lambda_t(\gamma_n(\rho_0))$ a.s. on $\{L = 0\}$. By assumption, $\gamma_n(\rho_0)$ is exchangeable, and Lemma 6.5 implies that $\gamma_n(\rho_t)$ is exchangeable. The assertion follows as the distribution of ρ_t is determined by the finite dimensional distributions. \square

For the application of Theorem 4.2, it is of interest whether the states ρ_t are a.s. dust-free. Proposition 6.6 formulates the criterion from [44, Proposition 30] in our present context. We call the finite measure Ξ on Δ dust-free if

$$\Xi\{0\} > 0 \quad \text{or} \quad \int |x|_1 |x|_2^{-2} \Xi_0(dx) = \infty. \quad (6.4)$$

Proposition 6.6. *Let $t \in (0, \infty)$ and assume $\rho_0 \in \mathfrak{U}$. Then Ξ is dust-free if and only if ρ_t is a.s. dust-free.*

Proof. By Remark 5.2, $\rho_t \in \mathfrak{U}$, hence $\Upsilon(\rho)$ is well-defined. Clearly, ρ_t is dust-free if and only if the partition $\Pi_s^{(t)}$ from Remark 6.4 contains no singletons for all $s \in (0, t) \cap \mathbb{Q}$. This holds a.s. if and only if Ξ is dust-free by [44, Proposition 30]. \square

7 Decomposition of the genealogical distances

To apply Theorem 4.2(i), we need to describe also the $\hat{\mathfrak{U}}$ -valued process $(\beta(\rho_t), t \in \mathbb{R}_+)$ by a martingale problem. A version of this process that readily yields a description by a martingale problem is read off from the lookdown model in this section. We define this process in Subsection 7.1 for a deterministic point measure η that drives the population model. In Subsection 7.2, we let η again be the Poisson random measure.

7.1 The deterministic construction

Let η be a simple point measure on $(0, \infty) \times \mathcal{P}$ as in Section 5. Let $(r_0, v_0) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$. We interpret (r_0, v_0) as a decomposition of genealogical distances at time 0. For $i \in \mathbb{N}$, let

$$\mathcal{P}(i) = \{\pi \in \mathcal{P} : \{i\} \notin \pi\}$$

be the set of partitions of \mathbb{N} in which i does not form a singleton block. If $\eta(\{s\} \times \mathcal{P}(A_s(t, i))) > 0$ for some $s \in (0, t]$, then we set

$$v_t(i) = t - \sup\{s \in (0, t] : \eta(\{s\} \times \mathcal{P}(A_s(t, i))) > 0\},$$

else we set

$$v_t(i) = t + v_0(A_0(t, i)).$$

The quantity $v_t(i)$ is the time back until an ancestor of the particle on level i at time t is involved in a reproduction event in which it belongs to a non-singleton block, if there is such an event, else $v_t(i)$ is defined from v_0 .

We let $\rho_0 = \alpha(r_0, v_0)$ and define the process $(\rho_t, t \in \mathbb{R}_+)$ from η and ρ_0 as in Section 5. We set

$$r_t(i, j) = (\rho_t(i, j) - v_t(i) - v_t(j)) \mathbf{1}\{i \neq j\}$$

for $t \in \mathbb{R}_+$ and $i, j \in \mathbb{N}$. Then (r_0, v_0) can be thought of as a decomposition of the distance matrix ρ_t in the sense of Section 2.

Remark 7.1. Consider the following change (compared to the beginning of Section 5) in the definition of the reproduction event encoded by a point $(t, \pi) \in \eta$: For each non-singleton block $B_i(\pi)$, the reproducing particle on level i at time $t-$ dies and is replaced at time t by its offspring on all the levels in $B_i(\pi)$. Then the quantity $v_t(i)$ is the age of the particle on level i at time t if this holds for $t = 0$. Condition (7.2) below ensures that the times at which the particles on a fixed level are replaced do not accumulate.

Analogously to Section 5, we give another description of the process $((r_t, v_t), t \in \mathbb{R}_+)$. Let \mathcal{S}_n be the set of semi-partitions of $[n]$, that is, the set of systems of nonempty disjoint subsets of $[n]$. Every partition is also a semi-partition. However, in a semi-partition, there can be missing elements, that is, elements of $[n]$ that are not contained in the union $\cup \sigma$ of the blocks of σ . By “blocks” we mean the subsets of $[n]$ that are the elements of σ . From every semi-partition $\sigma \in \mathcal{S}_n$, a partition π is obtained by inserting a singleton block for each missing element. We call π the partition associated with σ , and we define $\sigma(i) = \pi(i)$ for each $i \in [n]$, where $\pi(i)$ is defined in the beginning of Section 5. In order that equation (7.1) below hold, we associate with each element σ of \mathcal{S}_n a transformation $\mathbb{R}^{n^2} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n^2} \times \mathbb{R}^n$, which we also denote by σ , by $\sigma(r, v) = (r', v')$, where

$$v'(i) = v(\sigma(i)) \mathbf{1}\{i \notin \cup \sigma\}$$

and

$$r'(i, j) = (v(\sigma(i)) \mathbf{1}\{i \in \cup \sigma\} + r(\sigma(i), \sigma(j)) + v(\sigma(j)) \mathbf{1}\{j \in \cup \sigma\}) \mathbf{1}\{i \neq j\}$$

for $i, j \in [n]$.

We define the function

$$\varsigma_n : \mathcal{P} \rightarrow \mathcal{S}_n, \quad \pi \mapsto \{B \cap [n] : B \in \pi, \#B \geq 2\} \setminus \{\emptyset\}.$$

that removes all singleton blocks from a partition of \mathbb{N} and restricts the semi-partition obtained in this way to a semi-partition of $[n]$. The semi-partition $\varsigma_n(\pi)$ captures the effect on $\gamma_n(r_{t-}, v_{t-})$ of a reproduction event $(t, \pi) \in \eta$:

$$\varsigma_n(\pi)(\gamma_n(r_{t-}, v_{t-})) = \gamma_n(r_t, v_t) \tag{7.1}$$

Here we cannot use the restriction $\gamma_n(\pi)$ (of π to $[n]$) instead of $\varsigma_n(\pi)$ as we cannot read off from $\gamma_n(\pi)$ which singleton blocks in $\gamma_n(\pi)$ are also singleton blocks in π .

We define the set of partitions

$$\hat{\mathcal{P}}^n = \{\pi \in \mathcal{P} : \varsigma_n(\pi) = \emptyset\}.$$

We remark that $\hat{\mathcal{P}}^n$ is the set of partitions of \mathbb{N} in which not all of the first n integers form singleton blocks, hence it is strictly larger than the set \mathcal{P}^n . Only reproduction events that are encoded by a partition in $\hat{\mathcal{P}}^n$ affect the decomposed genealogical distances on the first n levels $(\gamma_n(r_t, v_t), t \in \mathbb{R}_+)$. If η satisfies the condition

$$\eta((0, t] \times \hat{\mathcal{P}}^n) < \infty \quad \text{for all } t \in (0, \infty) \text{ and } n \in \mathbb{N}. \quad (7.2)$$

then there are only finitely many reproduction events in bounded time intervals that result in a jump of the process $(\gamma_n(r_t, v_t), t \in \mathbb{R}_+)$. Between such jumps, the matrix r_t is constant, and the entries of the vector v_t grow linearly with slope 1, that is, $v_t(i) + s = v_{t+s}(i)$ for $i \in [n]$ and $t, s \in \mathbb{R}_+$ with $\eta((t, t+s] \times \hat{\mathcal{P}}^n) = 0$.

7.2 Stochastic evolution

Now let η be defined as the Poisson random measure from Section 6 whose distribution is characterized by some finite measure Ξ on Δ . Consider the population model from Section 5 driven by the Poisson random measure η , with the initial state defined as a $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ -valued random variable (r_0, v_0) that is independent of η . For each $n \in \mathbb{N}$ and $\sigma \in \mathcal{S}_n \setminus \{\emptyset\}$, the rate at which reproduction events encoded by a partition in $\varsigma_n^{-1}(\sigma) \in \mathcal{P}$ occur is given by

$$\begin{aligned} \lambda_{n,\sigma} &= H_{\Xi}(\varsigma_n^{-1}(\sigma)) \\ &= \int_{\Delta} \kappa(x, \varsigma_n^{-1}(\sigma)) |x|_2^{-2} \Xi_0(dx) + \Xi\{0\} \sum_{1 \leq i < j} \mathbf{1}\{K_{i,j} \in \varsigma_n^{-1}(\sigma)\} \\ &= \int_{\Delta} \sum_{\substack{i_1, \dots, i_{\ell} \in \mathbb{N} \\ \text{pairwise distinct}}} x_{i_1}^{k_1} \cdots x_{i_{\ell}}^{k_{\ell}} (1 - |x|_1)^{n-k_1-\dots-k_{\ell}} |x|_2^{-2} \Xi_0(dx) \\ &\quad + \Xi\{0\} \mathbf{1}\{\ell = 1, k_1 = 2\} + \infty \mathbf{1}\{\Xi\{0\} > 0, \ell = 1, k_1 = 1\} \end{aligned} \quad (7.3)$$

where $\ell = \#\sigma$, and $k_1, \dots, k_{\ell} \geq 1$ are the sizes of the subsets in σ in arbitrary order, and Ξ_0 is defined as in (6.1). For the last equality, we consider the paintbox partition π associated with $x \in \Delta$: With the notation from the beginning of Section 6, integers $i, j \in [n]$ are elements of a common subset in $\varsigma_n(\pi)$ if and only if U_i and U_j fall into a common subinterval that is not the dust interval. In particular, $i \notin \cup \varsigma_n(\pi)$ if and only if U_i falls into the dust interval.

Note that the rates λ_{π} for $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$, which we discussed already in Remark 6.1, satisfy

$$\lambda_{\pi} = H_{\Xi}(\gamma_n^{-1}(\pi)) = H_{\Xi}\left(\bigcup_{\sigma} \{\varsigma_n^{-1}(\sigma)\}\right) = \sum_{\sigma} \lambda_{n,\sigma}, \quad (7.4)$$

where the union and the sum are over all semi-partitions $\sigma \in \mathcal{S}_n$ with the same non-singleton blocks as π . In (7.4), we also use the restriction map $\gamma_n : \mathcal{P} \rightarrow \mathcal{P}_n$. From

equations (7.3) and (7.4), we see that $\lambda_{\{\{1,2\}\}} = \Xi(\Delta) < \infty$ and $\lambda_\pi < \infty$ for all $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$, where $\mathbf{0}_n = \{\{1\}, \dots, \{n\}\}$. This implies $\eta((0, t] \times \mathcal{P}^n) < \infty$ a.s. for all $t \in (0, \infty)$. That is, condition (5.1) is a.s. satisfied, as asserted in Section 6. The condition (6.4) for Ξ to be dust-free is the condition that $\lambda_{1, \{\{1\}\}} = \infty$. That is, each particle reproduces with infinite rate if and only if Ξ is dust-free. Hence, if Ξ is not dust-free, then almost every realization of η satisfies condition (7.2). Moreover, if Ξ is not dust-free, then $\lambda_{n, \sigma} < \infty$ for all $n \in \mathbb{N}$ and $\sigma \in \mathcal{S}_n \setminus \{\emptyset\}$ as a consequence of equation (7.3).

Remark 7.2. Consider the case that Ξ is concentrated on $\{(x, 0, 0, \dots) : x \in [0, 1]\} \subset \Delta$. In this case, which corresponds to the Λ -coalescent, a.s. no simultaneous multiple reproduction events occur. The measure Ξ_0 is then determined by the finite measure $\Lambda_0 = \varpi(\Xi_0)$, where $\varpi : \Delta \rightarrow [0, 1]$, $x \mapsto x_1$. For $B \subset [n]$ and $k = \#B$, it then follows

$$\lambda_{n, \{B\}} = \int_{[0, 1]} x^k (1-x)^{n-k} x^{-2} \Lambda_0(dx) + \Xi\{0\} \mathbf{1}\{k=2\} + \infty \mathbf{1}\{\Xi\{0\} > 0, k=1\}.$$

The rates $\lambda_{n, \sigma}$ for $\sigma \in \mathcal{S}_n$ with $\#\sigma > 1$ are equal to zero in this case.

Now we consider the process $((r_t, v_t), t \in \mathbb{R}_+)$ from Subsection 7.1, driven by the Poisson random measure η . The initial state is defined as a $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ -valued random variable (r_0, v_0) that is independent of η . Recall the set $\hat{\mathcal{C}}$ from Section 4. For $(r, v) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ and $\phi \in \hat{\mathcal{C}}_n$, we write

$$\langle \nabla^v \phi, \underline{1} \rangle(r, v) = \sum_{i \in \mathbb{N}} \frac{\partial}{\partial v(i)} \phi(r, v).$$

From the discussion above and the description of the process $(\gamma_n(r_t, v_t), t \in \mathbb{R}_+)$ in Section 7.1, we deduce the next proposition.

Proposition 7.3. *Assume that Ξ is not dust-free. Define an operator $\hat{A} = \hat{A}_{\text{grow}} + \hat{A}_{\text{repr}}$ with domain $\hat{\mathcal{C}}$ by*

$$\hat{A}_{\text{grow}} \phi(r, v) = \langle \nabla^v \phi, \underline{1} \rangle(r, v)$$

and

$$\hat{A}_{\text{repr}} \phi(r, v) = \sum_{\sigma \in \mathcal{S}_n \setminus \{\emptyset\}} \lambda_{n, \sigma} (\phi(\sigma(\gamma_n(r, v))) - \phi(r, v))$$

for $n \in \mathbb{N}$, $\phi \in \hat{\mathcal{C}}_n$ and $(r, v) \in \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$. Then the stochastic process $((r_t, v_t), t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{A}, \hat{\mathcal{C}})$.

The operator \hat{A}_{grow} accounts for the growth of the marks v_t which is described in the end of Subsection 7.1. The operator \hat{A}_{repr} stands for the jumps of the decomposed genealogical distances in reproduction events which are described by equation (7.1).

Let now $(\rho_t, t \in \mathbb{R}_+)$ be the process defined from η and some \mathfrak{U} -valued random variable ρ_0 that is independent of η . Then $\rho_t \in \mathfrak{U}$ for all $t \in \mathbb{R}_+$ by Remark 5.2. Assume

$$(r_0, v_0) = \beta(\rho_0). \tag{7.5}$$

Then the construction in Section 7.1 and the definition of the map α in Section 2 yield $(r_t, v_t) \in \hat{\mathfrak{U}}$ and $\rho_t = \alpha(r_t, v_t)$ for all $t \in \mathbb{R}_+$. The following proposition states that the decomposition (r_t, v_t) of the semi-ultrametric ρ_t is the one given by the map β from Section 2.

Proposition 7.4. *Assumption (7.5) implies $(r_t, v_t) = \beta(\rho_t)$ a. s. for each $t \in \mathbb{R}_+$.*

Corollary 7.5. *The process $(\beta(\rho_t), t \in \mathbb{R}_+)$ solves the martingale problem (\hat{A}, \hat{C}) .*

Proof. This is immediate from Propositions 7.3 and 7.4. \square

To prove Proposition 7.4, we use the following lemma.

Lemma 7.6. *Assume that Ξ is not dust-free. Let $t \in (0, \infty)$ and $i \in \mathbb{N}$. Then a. s. on the event $\{v_t(i) < t\}$, there exists an integer $k \in \mathbb{N} \setminus \{i\}$ with $v_t(i) = \frac{1}{2}\rho_t(i, k)$.*

Proof. Recall the process $(\Pi_s^{(t)}, s \in \mathbb{R}_+)$ from Remark 6.4. We work on the event of probability 1 on which $v_t(i) > 0$, condition (7.2) is satisfied, and for each $s \in (0, t) \cap \mathbb{Q}$, the partition $\Pi_s^{(t)}$ contains infinitely many blocks if it contains singletons. This is indeed an event of probability 1 as $\lambda_{1, \{\{1\}\}} < \infty$ and by Kingman's correspondence.

At time $t - v_i(t)$, a reproduction event occurs that is encoded by a partition in which the block that contains $A_{t-v_t(i)}(t, i)$ contains some other element j . This follows from the definition of $v_t(i)$ in Section 7.1 and as $\eta((0, t] \times \hat{\mathcal{P}}^i) < \infty$ by condition (7.2) which means that the reproduction events in which particles on levels not larger than i reproduce do not accumulate.

Moreover, by condition (7.2), there exists a time $s \in (t - v_i(t), t) \cap \mathbb{Q}$ with $\eta((t - v_t(i), s] \times \hat{\mathcal{P}}_j) = 0$, which implies that the particle on level j at time $t - v_t(i)$ is still on level j at time s .

By definition of $v_t(i)$, the partition $\Pi_s^{(t)}$ contains the singleton block $\{A_s(t, i)\}$, hence $\Pi_s^{(t)}$ has infinitely many blocks. This means that infinitely many particles at time s survive until time t . Remark 5.1 implies that all particles at time s survive until time t . Therefore, the particle that was on level j at the times $t - v_t(i)$ and s is on some level k at time t that satisfies $\frac{1}{2}\rho_t(i, k) = v_t(i)$. \square

Proof of Proposition 7.4. Let $t \in (0, \infty)$. From the definitions of the reproduction events in Section 5 and of the quantity $v_t(i)$ in Section 7.1, it follows that for each $s \in (t - v_t(i) \wedge t, t]$, only the particle on level i at time t descends from the particle on level $A_s(t, i)$ at time s . The definitions Υ in Section 2 and of ρ_t in Section 5 imply $0 \leq v_t(i) \wedge t \leq \Upsilon(\rho_t)(i) \wedge t$ for all $i \in \mathbb{N}$.

In the case that Ξ is dust-free, the assertion follows as $\Upsilon(\rho_t) = 0$ a. s. by Proposition 6.6.

Now we assume that Ξ is not dust-free. Let $i \in \mathbb{N}$. Lemma 7.6 yields $\Upsilon(\rho_t)(i) \leq v_t(i)$ a. s. on the event $\{v_t(i) < t\}$. On $\{v_t(i) \geq t\}$, the exchangeable partition $\Pi_t^{(t)}$, defined in Remark 6.4, contains the singleton block $\{A_0(t, i)\}$, and it follows by Kingman's correspondence that it has infinitely many blocks a. s. on $\{v_t(i) \geq t\}$. A. s. on $\{v_t(i) \geq t\}$, all particles at time zero survive until time t by Remark 5.1. Hence, as $\Upsilon(\rho_0) = v_0$ by assumption (7.5),

$$\begin{aligned} v_t(i) &= t + v_0(A_0(t, i)) = t + \frac{1}{2} \inf_{j \in \mathbb{N} \setminus \{A_0(t, i)\}} \rho_0(A_0(t, i), j) \\ &= \frac{1}{2} \inf_{j \in \mathbb{N} \setminus \{i\}} \rho_t(i, j) = \Upsilon(\rho_t)(i) \quad \text{a. s. on } \{v_t(i) \geq t\}. \end{aligned}$$

\square

8 Tree-valued Fleming-Viot processes

In this section, we apply Theorem 4.2 to the process $(\rho_t, t \in \mathbb{R}_+)$ from Section 6. We call all the image processes in Theorem 4.2 tree-valued Fleming-Viot processes. To distinguish them, we also call them \mathbb{U} -, $\hat{\mathbb{U}}$ -, and \mathcal{U}^{erg} -valued Ξ -Fleming-Viot processes. In this section, we also show uniqueness of the martingale problems for tree-valued Fleming-Viot processes.

8.1 Processes with values in the space of metric measure spaces

In this subsection, we consider a finite measure Ξ on Δ that is dust-free. Let $\chi \in \mathbb{U}$, and let $(\rho_t, t \in \mathbb{R}_+)$ be the \mathfrak{U} -valued Markov process from Section 6 that is defined in terms of Ξ and an initial state ρ_0 with distribution ν^χ . We define a \mathbb{U} -valued Ξ -Fleming-Viot process $(\chi_t, t \in \mathbb{R}_+)$ with initial state $\chi \in \mathbb{U}$ by $\chi_t = \psi(\rho_t)$. Remark 3.14 ensures that $\psi(\rho_0) = \chi$ a.s. By Theorem 4.2 and Propositions 6.2, 6.3, and 6.6, the process $(\chi_t, t \in \mathbb{R}_+)$ is Markovian and solves the martingale problem (B, Π) , where the generator B is defined by $B\Phi(\chi) = \nu^\chi(A\phi)$ for $\phi \in \mathcal{C}$ with associated polynomial $\Phi \in \Pi$, and $\chi \in \mathbb{U}$. Here A is the generator defined in Proposition 6.2. The martingale problem (B, Π) is a generalization of the martingale problem in Theorem 1 of Greven, Pfaffelhuber, and Winter [25].

Proposition 8.1. *The martingale problem (B, Π) is unique.*

Here uniqueness means that the finite-dimensional distributions of its solutions are uniquely specified when the initial state is given.

Proof. We use a function-valued dual process. This method is applied in the context of tree-valued Fleming-Viot processes in [13], another dual process is used in [25]. We fix $n \in \mathbb{N}$ and work with a dual process with state space \mathcal{C}_n . With each element π of \mathcal{P}_n , we also associate a transformation $\mathcal{C}_n \rightarrow \mathcal{C}_n$, which we also denote by π , by

$$\pi(\phi)(\rho) = \phi(\pi(\rho)).$$

We define an independent process $(\phi_t, t \in \mathbb{R}_+)$ as the Markov process with càdlàg paths in \mathcal{C}_n such that

- for each $\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}$ at rate λ_π , the process jumps from ϕ to $\pi(\phi)$,
- and between jumps, the process evolves deterministically according to

$$\phi_{t+s}(\rho) = \phi_t(\rho + \underline{\underline{2}}_n s)$$

for $s, t \in \mathbb{R}_+$ and $\rho \in \mathbb{R}^{n^2}$, where $\underline{\underline{2}}_n = 2(\mathbf{1}\{i \neq j\})_{i,j \in [n]}$.

The process $(\phi_t, t \in \mathbb{R}_+)$ solves the martingale problem $(B^\downarrow, \mathcal{D})$, where

$$\mathcal{D} = \{\mathcal{C}_n \rightarrow \mathbb{R}, \phi \mapsto \nu^{\chi'} \phi : \chi' \in \mathbb{U}\}$$

and an operator B^\downarrow with domain \mathcal{D} is defined by $B^\downarrow = B_{\text{coal}}^\downarrow + B_{\text{shrink}}^\downarrow$,

$$B_{\text{coal}}^\downarrow \nu^{\chi'}(\phi) = \sum_{\pi \in \mathcal{P}_n \setminus \{\mathbf{0}_n\}} \lambda_\pi \left(\nu^{\chi'}(\pi(\phi)) - \nu^{\chi'} \phi \right)$$

and

$$B_{\text{shrink}}^\downarrow \nu^{\chi'}(\phi) = \nu^{\chi'} \langle \nabla \phi, \underline{\underline{2}} \rangle$$

for $\phi \in \mathcal{C}_n$ and $\chi' \in \mathbb{U}$.

From this definition, we have $B(\nu\phi)(\chi') = B^\downarrow \nu^{\chi'}(\phi)$ for all $\phi \in \mathcal{C}_n$ and $\chi' \in \mathbb{U}$, where $\nu\phi$ is the polynomial associated with ϕ . For all $t \in \mathbb{R}_+$ and all polynomials $\Phi \in \Pi$ of degree at most n , it follows from Theorem 4.4.11 in [18] that $E[\Phi(\tilde{\chi}_t)]$ is equal for all solutions $((\tilde{\chi}_t, t \in \mathbb{R}_+); P)$ of the martingale problem (B, Π) with initial state χ_0 . As $n \in \mathbb{N}$ was arbitrary and the space Π of polynomials is separating, the uniqueness assertion follows from Theorem 4.4.2 in [18]. \square

8.2 Processes with values in the space of marked metric measure spaces

Let Ξ be a general finite measure on the simplex Δ . Let $\hat{\chi} \in \hat{\mathbb{U}}$, let ρ_0 be a \mathfrak{U} -valued random variable with distribution $\alpha(\nu^{\hat{\chi}})$, and let the \mathfrak{U} -valued Markov process $(\rho_t, t \in \mathbb{R}_+)$ be defined, as in Section 6, from Ξ and the initial state ρ_0 . We define a $\hat{\mathbb{U}}$ -valued Ξ -Fleming-Viot process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ with initial state $\chi \in \hat{\mathbb{U}}$ by $\hat{\chi}_t = \hat{\psi}(\beta(\rho_t))$ for $t \in \mathbb{R}_+$. By Propositions 3.3 and 3.12, the initial state satisfies $\hat{\chi}_0 = \chi$ a.s.

If Ξ is not dust-free, then by Theorem 4.2 and Propositions 6.3 and 7.3, the process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ is Markovian and solves the martingale problem $(\hat{B}, \hat{\Pi})$, where the generator \hat{B} is defined by $\hat{B}\Phi(\chi') = \nu^{\chi'}(\hat{A}\phi)$ for all $\phi \in \hat{\mathcal{C}}$ with associated marked polynomial Φ , and all $\chi' \in \hat{\mathbb{U}}$. Here the generator \hat{A} is defined as in Proposition 7.3. Also the martingale problem $(\hat{B}, \hat{\Pi})$ is unique, the proof is analogous to Proposition 8.1.

If Ξ is dust-free, then for each $t \in \mathbb{R}_+$, the marked metric measure space $\hat{\chi}_t$ is a.s. dust-free, hence $\hat{\chi}_t$ is given a.s. by the associated metric measure space χ_t , cf. Remark 4.4. In this case, the process $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves a martingale problem that is similar to (B, Π) .

8.3 Processes with values in the space of distance matrix distributions

Let $(\hat{\chi}_t, t \in \mathbb{R}_+)$ be the process from Section 8.2, where Ξ is a general finite measure on the simplex Δ . We define a \mathcal{U}^{erg} -valued Ξ -Fleming-Viot process $(\nu_t, t \in \mathbb{R}_+)$ with initial state $\alpha(\nu^{\hat{\chi}_0}) \in \mathcal{U}^{\text{erg}}$ by $\nu_t = \alpha(\nu^{\hat{\chi}_t})$. Again by Theorem 4.2 and Propositions 6.2 and 6.3, it follows that $(\nu_t, t \in \mathbb{R}_+)$ is Markovian and solves the martingale problem (C, \mathcal{C}) , where the generator C is defined by $C\Psi(\nu) = \nu(A\phi)$ for all $\nu \in \mathcal{U}^{\text{erg}}$, $\phi \in \mathcal{C}$, and $\Psi \in \mathcal{C} : \nu' \mapsto \nu'\phi$. Here the generator A is defined as in Proposition 6.2. Uniqueness of the martingale problem follows analogously to Proposition 8.1, we use the same \mathcal{C}_n -valued dual process.

Remark 8.2. The martingale problems characterize only versions of the processes in Subsections 8.1 – 8.3. Moreover, Proposition 3.16 shows $\beta(\rho_t) \in \hat{\mathfrak{D}}^*$ (and in the dust-free case also $\rho_t \in \mathfrak{D}^*$ by Corollary 3.17) only for a fixed t (or countably many t) on an event of probability one. At the other times, we do not exclude in the present article that $\hat{\psi}(\beta(\rho_t))$ is just the arbitrary state in the definition of $\hat{\psi}$. By techniques specific to the lookdown model, it is shown in [27] that the aforementioned assertions on ρ_t also hold

simultaneously for all $t \in \mathbb{R}_+$ on an event of probability one (see Theorems 3.1(i) and 3.9(i), and Remarks 4.4 and 4.13 in [27]).

9 Some semigroup properties

In this section, we use the lookdown construction to study Feller continuity of tree-valued Ξ -Fleming-Viot processes, and to show that the domains of the martingale problems for them are cores. We consider $\hat{\mathcal{U}}$ -valued Ξ -Fleming-Viot processes in detail, analogous results hold for the other processes from Section 8.

Let Ξ be a finite measure on the simplex Δ . For $\chi \in \hat{\mathcal{U}}$, let $(\hat{\chi}_t, t \in \mathbb{R}_+)$ under the probability measure \mathbb{P}_χ with associated expectation \mathbb{E}_χ be the $\hat{\mathcal{U}}$ -valued Ξ -Fleming-Viot process from Section 8.2 with initial state χ . We denote by $C_b(E)$ the set of bounded continuous \mathbb{R} -valued functions on a metric space E . We endow $C_b(E)$ with the supremum norm.

Proposition 9.1 states the Feller continuity of a $\hat{\mathcal{U}}$ -valued Ξ -Fleming-Viot process, namely that its semigroup preserves the set of bounded continuous functions.

Proposition 9.1. *For each $t \in \mathbb{R}_+$ and $f \in C_b(\hat{\mathcal{U}})$, the map $\hat{\mathcal{U}} \rightarrow \mathbb{R}$, $\chi \mapsto \mathbb{E}_\chi[f(\hat{\chi}_t)]$ is continuous.*

Proof. We fix $t \in \mathbb{R}_+$, and we denote by \mathcal{N} the space of simple point measures on $(0, \infty) \times \mathcal{P}$. Recall the deterministic construction from Section 7.1 and denote by $g : \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}} \times \mathcal{N} \rightarrow \mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$ the function that maps the initial state (r_0, v_0) and the simple point measure η to (r_t, v_t) . Clearly, the function g is continuous in $\mathbb{R}^{\mathbb{N}^2} \times \mathbb{R}^{\mathbb{N}}$. Let $\phi \in \hat{\mathcal{C}}_n$. By definition of the marked Gromov-weak topology and dominated convergence, also the function

$$h : \hat{\mathcal{U}} \rightarrow \mathbb{R}, \quad \chi \mapsto \int \nu^\chi(d(r, v)) \int \mathbb{P}(\eta \in d\eta') \phi \circ g((r, v), \eta')$$

is continuous, where η is the Poisson random measure from Section 6.

For $\chi \in \hat{\mathcal{U}}$, under \mathbb{P}_χ , let (r_0, v_0) be a $\hat{\mathcal{U}}$ -valued random variable with distribution ν^χ . Let (r_t, v_t) be the $\hat{\mathcal{U}}$ -valued random variable defined from (r_0, v_0) and Ξ as in Section 7.2. By construction, $h(\chi) = \mathbb{E}_\chi[\phi(r_t, v_t)]$. Let $\Phi \in \hat{\Pi}$ be the marked polynomial associated with ϕ . By Propositions 7.4 and 4.7, and as we may assume $\hat{\psi}(r_t, v_t) = \hat{\chi}_t$ a.s., it follows that $h(\chi) = \mathbb{E}_\chi[\Phi(\hat{\chi}_t)]$, hence $\chi \mapsto \mathbb{E}_\chi[\Phi(\hat{\chi}_t)]$ is continuous for each marked polynomial $\Phi \in \hat{\Pi}$. The assertion follows as the set $\hat{\Pi}$ of marked polynomials is convergence determining and by definition of convergence in distribution in $\hat{\mathcal{U}}$. \square

Using the lookdown construction, we will also show the following lemma.

Lemma 9.2. *For each $t \in \mathbb{R}_+$ and $\Phi \in \hat{\Pi}$, the function $\hat{\mathcal{U}} \rightarrow \mathbb{R}$, $\chi \mapsto \mathbb{E}_\chi[\Phi(\hat{\chi}_t)]$ is an element of $\hat{\Pi}$.*

Let \hat{L} denote the closure of $\hat{\Pi}$ in $C_b(\hat{\mathcal{U}})$ with respect to the supremum norm. For application in [26], we note two corollaries of Lemma 9.2. The first corollary states that the semigroup of a $\hat{\mathcal{U}}$ -valued Ξ -Fleming-Viot process can be restricted to a semigroup on \hat{L} that is strongly continuous.

Corollary 9.3. *Let $f \in \hat{L}$. Then for each $t \in \mathbb{R}_+$, the function $\hat{\mathcal{U}} \rightarrow \mathbb{R}$, $\chi \mapsto \mathbb{E}_\chi[f(\hat{\chi}_t)]$ is an element of \hat{L} . Moreover,*

$$\limsup_{t \downarrow 0} \sup_{\chi \in \hat{\mathcal{U}}} |\mathbb{E}_\chi[f(\hat{\chi}_t)] - \mathbb{E}_\chi[f(\hat{\chi}_0)]| = 0.$$

Proof. The first assertion follows from Lemma 9.2 and the definition of \hat{L} . As $(\hat{\chi}_t, t \in \mathbb{R}_+)$ solves the martingale problem $(\hat{B}, \hat{\Pi})$ from Section 8.2,

$$\mathbb{E}_\chi[\Phi(\hat{\chi}_t)] - \mathbb{E}_\chi[\Phi(\hat{\chi}_0)] = \mathbb{E}_\chi\left[\int_0^t \hat{B}\Phi(\hat{\chi}_s)ds\right]$$

for all $t \in \mathbb{R}_+$ and $\Phi \in \hat{\Pi}$. The second assertion follows as $\hat{B}\Phi$ is bounded and by definition of \hat{L} . \square

The next corollary says that the semigroup on \hat{L} of a $\hat{\mathcal{U}}$ -valued Ξ -Fleming-Viot process is generated by the closure of the operator \hat{B} with domain $\hat{\Pi}$, see [18, Chapter 1] for the definitions.

Corollary 9.4. *The subspace $\hat{\Pi} \subset C_b(\hat{\mathcal{U}})$ is a core for the generator of the semigroup on \hat{L} of a $\hat{\mathcal{U}}$ -valued Ξ -Fleming-Viot process.*

Proof. We note that \hat{B} is the restriction of the generator of the semigroup to $\hat{\Pi}$ and apply Propositions 1.3.3 and Corollary 1.1.6 of [18], using Lemma 9.2 and Corollary 9.3. \square

Proof of Lemma 9.2. Let the space \mathcal{N} be defined as in the proof of Proposition 9.1. Let $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$. Note that in the construction in Sections 5 and 7.1, the restriction $\gamma_n(r_t, v_t)$ depends only on the simple point measure η and the restriction $\gamma_n(r_0, v_0)$ of the initial state. We may thus define the function $g_n : \mathbb{R}^{n^2} \times \mathbb{R}^n \times \mathcal{N} \rightarrow \mathbb{R}^{n^2} \times \mathbb{R}^n$ that maps the restriction $\gamma_n(r_0, v_0)$ of the initial state and the point measure η to $\gamma_n(r_t, v_t)$. Note that when the simple point measure is fixed, g_n is a differentiable function on $\mathbb{R}^{n^2} \times \mathbb{R}^n$ with bounded uniformly continuous derivative.

Let $\phi \in \hat{\mathcal{C}}_n$. We define the function

$$f : \mathbb{R}^{n^2} \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (r, v) \mapsto \int \mathbb{P}(\eta \in d\eta') \phi \circ g_n((r, v), \eta'),$$

where η is the Poisson random measure from Section 6. By dominated convergence and the mean value theorem, also the function f is differentiable with bounded uniformly continuous derivative, and we obtain that $f \in \hat{\mathcal{C}}_n$.

Let Φ be the marked polynomial associated with ϕ . As in the proof of Proposition 9.1, Propositions 7.4 and 4.7 imply $\mathbb{E}_\chi[\Phi(\hat{\chi}_t)] = \nu^\chi f$ for all $\chi \in \hat{\mathcal{U}}$. Hence, $\chi \mapsto \mathbb{E}_\chi[\Phi(\hat{\chi}_t)]$ is in $\hat{\Pi}$. \square

Let L be the closure of Π in $C_b(\mathcal{U})$ and let L' be the closure of \mathcal{C} in $C_b(\mathcal{U}^{\text{erg}})$, with respect to the supremum norm. In the same way as above, it can be shown: The semigroup on L' of a \mathcal{U}^{erg} -valued Ξ -Fleming-Viot process is strongly continuous and generated by the closure of the operator C with domain \mathcal{C} from Section 8.3. If Ξ is dust-free, then the semigroup on L of a \mathcal{U} -valued Ξ -Fleming-Viot process is strongly continuous and generated by the closure of the operator B with domain Π from Section 8.1. Continuity properties analogous to Proposition 9.1 also hold.

10 Convergence to equilibrium

Let Ξ be a finite measure on the simplex Δ with $\Xi(\Delta) > 0$. We show convergence to equilibrium for the $\hat{\mathcal{U}}$ -valued process $(\beta(\rho_t), t \in \mathbb{R}_+)$ from Section 7.2. From this, we deduce that also the tree-valued Ξ -Fleming-Viot process from Section 8.2 converges to equilibrium. In the same way, it can be shown that the other processes from Section 8 converge to equilibrium.

We define stationary processes and use a coupling argument. Analogously to Section 6, let $\bar{\eta}$ be a Poisson random measure on $\mathbb{R} \times \mathcal{P}$ with intensity $dt H_\Xi(d\pi)$. This Poisson random measure drives a population model in two-sided time (indexed by \mathbb{R}) where the reproduction events and the ancestral levels $\bar{A}_s(t, i)$ are defined as in Section 5. Then we define the stationary \mathcal{U} -valued process $(\bar{\rho}_t, t \in \mathbb{R})$ of the genealogical distances by

$$\bar{\rho}_t(i, j) = 2t - 2 \sup\{s \in (-\infty, t] : \bar{A}_s(t, i) = \bar{A}_s(t, j)\}$$

for $t \in \mathbb{R}$, $i, j \in \mathbb{N}$. On an event of probability 1, all these distances are finite. This follows from the assumption that $\Xi(\Delta) > 0$. That $\bar{\rho}_t$ is indeed a semi-ultrametric for each $t \in \mathbb{R}$ can be seen as in Remark 5.2. Clearly, ρ_t is exchangeable, which follows from exchangeability of the Ξ -coalescent as in Remark 6.4, or from Proposition 6.3.

Let η denote the restriction of $\bar{\eta}$ to $(0, \infty) \times \mathcal{P}$. Let $\chi \in \hat{\mathcal{U}}$ be arbitrary, and let ρ_0 be a \mathcal{U} -valued random variable with distribution $\alpha(\nu^\chi)$, independent of η . Let the process $(\rho_t, t \in \mathbb{R}_+)$ be defined from ρ_0 and η as in Section 5. By the construction in Section 5, on the event $\{\max_{i,j \in [n]} \bar{\rho}_t(i, j) < 2t\}$, the marked distance matrix $\gamma_n(\beta(\rho_t))$ does not depend on ρ_0 . As $\bar{\rho}_t$ can also be obtained from $\bar{\rho}_0$ and η as in Section 5, it follows that $\gamma_n(\beta(\rho_t)) = \gamma_n(\beta(\bar{\rho}_t))$ on the event $\{\max_{i,j \in [n]} \bar{\rho}_t(i, j) < 2t\}$. By stationarity, it follows that

$$|\mathbb{E}[\phi(\beta(\rho_t))] - \mathbb{E}[\phi(\beta(\bar{\rho}_0))]| \leq 2 \sup |\phi| \mathbb{P}(\max_{i,j \in [n]} \bar{\rho}_0(i, j) \geq 2t) \rightarrow 0 \quad (t \rightarrow \infty) \quad (10.1)$$

for all $\phi \in \hat{\mathcal{C}}_n$.

We call a $\hat{\mathcal{U}}$ -valued random variable that is distributed as $\bar{\chi}_0 := \hat{\psi}(\beta(\bar{\rho}_0))$ a Ξ -coalescent measure tree, generalizing the Λ -coalescent measure tree from [25]. A $\hat{\mathcal{U}}$ -valued Ξ -Fleming-Viot process $(\chi_t, t \in \mathbb{R}_+)$ with initial state χ is given by $\chi_t = \hat{\psi}(\beta(\rho_t))$, as in Section 8.2. As in Proposition 4.7, we obtain $\mathbb{E}[\Phi(\bar{\chi}_0)] = \mathbb{E}[\phi(\bar{\rho}_0)]$ and $\mathbb{E}[\Phi(\chi_t)] = \mathbb{E}[\phi(\rho_t)]$. The convergence (10.1) now yields that $\mathbb{E}[\Phi(\chi_t)]$ converges to $\mathbb{E}[\Phi(\bar{\chi}_0)]$ as $t \rightarrow \infty$ for all marked polynomials $\Phi \in \hat{\Pi}$. Using that the set $\hat{\Pi}$ is convergence determining in $\hat{\mathcal{U}}$, we deduce the following proposition.

Proposition 10.1. *The $\hat{\mathcal{U}}$ -valued random variable χ_t converges in distribution to a Ξ -coalescent measure tree as $t \rightarrow \infty$.*

A stationary $\hat{\mathcal{U}}$ -valued Ξ -Fleming-Viot process can be defined by $(\hat{\psi}(\beta(\bar{\rho}_t)), t \in \mathbb{R})$. In [25, Theorem 3], duality is used to show that the tree-valued Fleming-Viot process converges to an equilibrium. In [16, Theorem 4.1], convergence to stationarity of measure-valued Fleming-Viot processes is also proved by a coupling argument.

11 Proofs

In Section 11.1, we prove that the maps ψ and $\hat{\psi}$ from Section 3.4 are measurable. In Section 11.2, we prove Proposition 3.16 which is central for the proof of Theorem 1.2. We use the coupling characterization of the Prohorov metric, see e.g. [18, Theorem 3.1.2]: The Prohorov distance between two probability measures μ and μ' on the Borel sigma algebra on a separable metric space (Z, d^Z) is given by

$$d_P^Z(\mu, \mu') = \inf_{\nu} \inf \{ \varepsilon > 0 : \nu \{ (x, y) \in Z^2 : d^Z(x, y) > \varepsilon \} < \varepsilon \}, \quad (11.1)$$

where the first infimum is over all couplings ν of the probability measures μ and μ' .

11.1 Measurability of the construction of (marked) metric measure spaces

We show measurability of the map $\hat{\psi}$. Measurability of ψ follows along the same lines.

We use the marked Gromov-Prohorov metric d_{mGP} which metrizes the marked Gromov-weak topology on $\hat{\mathbb{M}}$, see [12]. It is defined by

$$d_{\text{mGP}}((X, r, m), (X', r', m')) = \inf_{Z, \varphi, \varphi'} d_P^Z(\hat{\varphi}(m), \hat{\varphi}'(m')).$$

Here the infimum is over all isometric embeddings $\varphi : X \rightarrow Z$ and $\varphi' : X' \rightarrow Z$ into complete and separable metric spaces (Z, d^Z) . The space $Z \times \mathbb{R}_+$ is endowed with the product metric $d^{Z \times \mathbb{R}_+}((z, v), (z', v')) = d^Z(z, z') \vee |v - v'|$. The maps $\hat{\varphi} : X \times \mathbb{R}_+ \rightarrow Z \times \mathbb{R}_+$ and $\hat{\varphi}' : X' \times \mathbb{R}_+ \rightarrow Z \times \mathbb{R}_+$ are defined by $\hat{\varphi}(x, v) = (\varphi(x), v)$, $(x, v) \in X \times \mathbb{R}_+$ and $\hat{\varphi}'(x', v) = (\varphi'(x'), v)$, $(x', v) \in X' \times \mathbb{R}_+$.

We write $\hat{\mathfrak{D}} = \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}}$. For $n \in \mathbb{N}$, we denote by

$$\begin{aligned} \hat{\mathfrak{D}}_n = \{ (r, v) \in \mathbb{R}_+^{n^2} \times \mathbb{R}_+^n : & r(i, i) = 0, r(i, j) = r(j, i), \\ & r(i, j) + r(j, k) \geq r(i, k) \text{ for all } i, j, k \in [n] \} \end{aligned}$$

the space of decomposed semimetrics on $[n]$ which we view as a subspace of $\mathbb{R}^{n^2} \times \mathbb{R}^n$. We denote by $\hat{\psi}_n : \hat{\mathfrak{D}}_n \rightarrow \hat{\mathbb{M}}$ the function that maps $(r, v) \in \hat{\mathfrak{D}}_n$ to the isomorphism class of the marked metric measure space $([n], r, n^{-1} \sum_{i=1}^n \delta_{(i, v(i))})$, here we also identify the elements of the semi-metric space $([n], r)$ with distance zero.

Remark 11.1. Clearly, the map $\hat{\psi}_n$ is continuous. For convenience, we state a proof. W.l.o.g. we can assume that $\hat{\mathfrak{D}}$ is endowed with the metric d that is given by

$$d((r, v), (r', v')) = \sup_{k \in \mathbb{N}} ((\max_{i, j \in [k]} |r(i, j) - r'(i, j)| \vee \max_{i \in [k]} |v(i) - v'(i)|) \wedge (2^{-k}))$$

for all $(r, v), (r', v') \in \hat{\mathfrak{D}}$. For $(r, v), (r', v') \in \hat{\mathfrak{D}}_n$, we define a probability measure ν on $(\hat{\mathfrak{D}})^2$ as the distribution of $((r(x_i, x_j))_{i, j \in \mathbb{N}}, (\tilde{v}_i)_{i \in \mathbb{N}}, (r'(x'_i, x'_j))_{i, j \in \mathbb{N}}, (\tilde{v}'_i)_{i \in \mathbb{N}})$, where $(x_i, \tilde{v}_i, x'_i, \tilde{v}'_i)_{i \in \mathbb{N}}$ is an iid sequence with distribution $n^{-1} \sum_{k=1}^n \delta_{(k, v(k), k, v'(k))}$. Then $\nu(\cdot \times \hat{\mathfrak{D}}) = \nu^{\hat{\psi}_n(r, v)}$ and $\nu(\hat{\mathfrak{D}} \times \cdot) = \nu^{\hat{\psi}_n(r', v')}$. For

$$c := \max_{i, j \in [n]} |r(i, j) - r'(i, j)| \vee \max_{i \in [n]} |v(i) - v'(i)|,$$

the coupling characterization (11.1) implies

$$d_P(\nu^{\hat{\psi}_n(r,v)}, \nu^{\hat{\psi}_n(r',v')}) \leq c + \nu\{(y, y') \in \hat{\mathfrak{D}}^2 : d(y, y') \geq c\} = c.$$

Continuity of $\hat{\psi}_n$ follows by definition of the marked Gromov-weak topology.

Proof of Proposition 3.8. Let $(r, v) \in \hat{\mathfrak{D}}^*$ and let (X, r) be the metric completion of (\mathbb{N}, r) . We endow the product space $X \times \mathbb{R}_+$ with the metric $d^{X \times \mathbb{R}_+}((x, v), (x', v')) = r(x, v) \vee |v - v'|$. The definition of $\hat{\mathfrak{D}}^*$ yields $\lim_{n \rightarrow \infty} d_P^{X \times \mathbb{R}_+}(n^{-1} \sum_{i=1}^n \delta_{(i, v(i))}, m) = 0$ for a probability measure m on $X \times \mathbb{R}_+$. As $\hat{\psi}(r, v)$ equals the isomorphism class of (X, r, m) , and as $\hat{\psi}_n(\gamma_n(r, v))$ equals the isomorphism class of $(X, r, n^{-1} \sum_{i=1}^n \delta_{(i, v(i))})$ for each $n \in \mathbb{N}$, the definition of the marked Gromov-Prohorov metric implies that $\lim_{n \rightarrow \infty} d_{\text{mGP}}(\hat{\psi}(r, v), \hat{\psi}_n(\gamma_n(r, v))) = 0$. Using Remark 11.1 and that $\hat{\mathfrak{D}}^*$ is a measurable subset of $\hat{\mathfrak{D}}$, we deduce measurability of $\hat{\psi}$. \square

11.2 Construction of the marked metric measure space in the sampling representation

We give three proofs of Proposition 3.16 that build on a common part, namely statement (11.4) below. The plan for the first proof is the following: We partition the completion of the tree (T, d) associated with the semi-ultrametric ρ (as in Remark 1.1) into small subsets. Into each of these subsets, we lay an atom whose mass is given by the asymptotic frequency of those integers that label the leaves of T that are the endpoints of the external branches that begin in this subset. By exchangeability, these asymptotic frequencies exist, and (11.4) yields that they add up to one. We obtain an atomic probability measure on the product space of the metric completion of the tree and the mark space \mathbb{R}_+ by defining the \mathbb{R}_+ -component as the distance to the top of the coalescent tree. Using the coupling characterization of the Prohorov metric, we show that this probability measure converges as the subsets become infinitely small, and that the limit measure coincides with the limit of the uniform measures in the definition of $\hat{\mathfrak{D}}^*$.

As a slight difference to the description in the preceding paragraph, we will work with the space (X, r) that corresponds to the completion of the space only of the starting vertices of the external branches, but we will occasionally recall the relation to the whole tree. We will use definitions from Section 2.

Proof of Proposition 3.16. Let $(r, v) = \beta(\rho)$. Then $v = \Upsilon(\rho)$ by definition of the map β . Let (X, r) be the metric completion of the semi-metric space (\mathbb{N}, r) .

Let $\varepsilon > 0$. As the distribution of the random variable $v(i)$ has at most countably many atoms, there exists a deterministic sequence $0 < h_1^{(\varepsilon)} < h_2^{(\varepsilon)} < \dots$ that increases to infinity and that satisfies

$$h_1^{(\varepsilon)} < \varepsilon, \quad h_{n+1}^{(\varepsilon)} - h_n^{(\varepsilon)} < \varepsilon,$$

and

$$\mathbb{P}(v(i) = h_n^{(\varepsilon)}) = 0 \tag{11.2}$$

for all $i, j, n \in \mathbb{N}$. We set $h_0^{(\varepsilon)} = 0$ and we write $I_n^\varepsilon = [h_{n-1}^{(\varepsilon)}, h_n^{(\varepsilon)})$ for $n \in \mathbb{N}$.

We define an equivalence relation \sim^ε on \mathbb{N} such that two distinct integers i, j are equivalent if and only if there exists $n \in \mathbb{N}$ with

$$v(i), v(j), \frac{1}{2}\rho(i, j) \in I_n^\varepsilon.$$

To show transitivity, we consider $i, j, k \in \mathbb{N}$ with $i \neq k$, $i \sim^\varepsilon j$, and $j \sim^\varepsilon k$. Then there exists $n \in \mathbb{N}$ with $v(i), v(j), v(k), \rho(i, j)/2, \rho(j, k)/2 \in I_n^\varepsilon$. As

$$v(i) \leq \rho(i, k)/2 \leq (\rho(i, j) \vee \rho(j, k))/2$$

by definition of Υ and ultrametricity, it follows that $i \sim^\varepsilon k$.

Note that the definitions in Section 2 imply

$$r(i, j) = (\frac{1}{2}\rho(i, j) - v(i) + \frac{1}{2}\rho(i, j) - v(j)) \mathbf{1}\{i \neq j\} < 2\varepsilon \quad (11.3)$$

for $i \sim^\varepsilon j$. (That is, in the context of Remark 2.2, the starting points of external branches that end in leaves $(0, i)$, $(0, j)$ of T with $i \sim^\varepsilon j$ have distance smaller than 2ε .)

In the next two paragraphs, we prove the following claim:

$$\text{A.s., the partition of } \mathbb{N} \text{ given by } \sim^\varepsilon \text{ contains no singleton blocks.} \quad (11.4)$$

For each $i, n \in \mathbb{N}$ the sequence $(\mathbf{1}\{v(j) \in I_n^\varepsilon, \rho(i, j)/2 \in I_n^\varepsilon\}, j \in \mathbb{N} \setminus \{i\})$ is exchangeable. By the de Finetti theorem, it is conditionally iid. Hence, on the event that there exists $j \in \mathbb{N} \setminus \{i\}$ with $v(j) \in I_n^\varepsilon$ and $\rho(i, j)/2 \in I_n^\varepsilon$, there exists a.s. another (in fact, infinitely many) such j in $\mathbb{N} \setminus \{i\}$.

For $j \in \mathbb{N}$, the definition of Υ and condition (11.2) imply the existence of (random) $n \in \mathbb{N}$ and $i \in \mathbb{N} \setminus \{j\}$ such that $v(j) \in I_n^\varepsilon$ and $\rho(i, j)/2 \in I_n^\varepsilon$ a.s. As shown in the preceding paragraph, there exists a.s. an integer $k \in \mathbb{N} \setminus \{i, j\}$ with $v(k) \in I_n^\varepsilon$ and $\rho(i, k)/2 \in I_n^\varepsilon$. From

$$v(k) \leq \rho(j, k)/2 \leq (\rho(i, j) \vee \rho(i, k))/2,$$

it follows that $\rho(j, k)/2 \in I_n^\varepsilon$ a.s. This proves (11.4).

Now we show that the asymptotic frequencies exist and add up to one. For $A \subset \mathbb{N}$ and $k \in \mathbb{N}$, we denote the relative frequency by $|A|_k = k^{-1}\#(A \cap [k])$ and the asymptotic frequency by $|A| = \lim_{k \rightarrow \infty} |A|_k$, provided the limit exists. As the random partition given by \sim^ε is exchangeable, the asymptotic frequencies of its blocks exist a.s. by Kingman's correspondence. Let $B^\varepsilon(i)$ denote the equivalence class of $i \in \mathbb{N}$ with respect to \sim^ε , and let

$$M^\varepsilon = \{j \in \mathbb{N} : j = \min B^\varepsilon(i) \text{ for some } i \in \mathbb{N}\}$$

be the set of minimal elements of the equivalence classes of \sim^ε . As the exchangeable partition given by \sim^ε has no singleton blocks a.s., it has proper frequencies by Kingman's correspondence, that is,

$$\sum_{i \in M^\varepsilon} |B^\varepsilon(i)| = 1 \quad \text{a.s.}$$

Consequently, on an event of probability 1, a probability measure m^ε on the product sigma algebra on $X \times \mathbb{R}_+$ is given by

$$m^\varepsilon = \sum_{i \in M^\varepsilon} |B^\varepsilon(i)| \delta_{(i, v(i))}. \quad (11.5)$$

(Into each of the subsets of (X, r) given by \sim^ε , the first component of the measure m^ε lays an atom with mass given by the asymptotic frequency of the integers that label the corresponding leaves in T .)

Let $\varepsilon_1 > \varepsilon_2 > \dots > 0$ with $\lim_{\ell \rightarrow \infty} \varepsilon_\ell = 0$. For each $\ell \in \mathbb{N}$, we replace ε with ε_ℓ everywhere in this proof until now, and we use the notations introduced so far. We also assume that for $k \leq \ell$, the sequence $(h_n^{(\varepsilon_k)}, n \in \mathbb{N})$ is contained in $(h_n^{(\varepsilon_\ell)}, n \in \mathbb{N})$. That is, the partitions $\{I_n^{\varepsilon_k}, n \in \mathbb{N}\}$ of \mathbb{R}_+ are nested.

For each $k \leq \ell$ and $i \in M^{\varepsilon_k}$, there exist a. s. $i_1, i_2, \dots \in M^{\varepsilon_\ell}$ with

$$B^{\varepsilon_k}(i) = B^{\varepsilon_\ell}(i_1) \uplus B^{\varepsilon_\ell}(i_2) \uplus \dots$$

By Fatou's lemma and as a. s., the partition given by \sim^{ε_ℓ} has proper frequencies, it follows that

$$|B^{\varepsilon_k}(i)| = |B^{\varepsilon_\ell}(i_1)| + |B^{\varepsilon_\ell}(i_2)| + \dots \quad \text{a. s.}$$

Hence, a coupling K of the probability measures m^{ε_k} and m^{ε_ℓ} on $X \times \mathbb{R}_+$ is given a. s. by

$$K\{((i, v(i)), (j, v(j)))\} = m^{\varepsilon_\ell}\{(j, v(j))\}$$

for $i \in M^{\varepsilon_k}$ and $j \in M^{\varepsilon_\ell}$ with $j \in B^{\varepsilon_k}(i)$. By construction, $i \sim^{\varepsilon_k} j$, hence $|v(i) - v(j)| < \varepsilon_k$ and $r(i, j) < 2\varepsilon_k$. By definition of m^ε and the coupling characterization of the Prohorov metric (11.1), this implies

$$d_P^{X \times \mathbb{R}_+}(m^{\varepsilon_k}, m^{\varepsilon_\ell}) \leq 2\varepsilon_k \quad (11.6)$$

a. s. for all $k \leq \ell$, when $X \times \mathbb{R}_+$ is endowed with the product metric $d^{X \times \mathbb{R}_+}$ that is given by $d^{X \times \mathbb{R}_+}((x, v), (x', v')) = r(x, x') \vee |v - v'|$. As a consequence, on an event of probability 1, the sequence $(m^{\varepsilon_\ell}, \ell \in \mathbb{N})$ in the space of probability measures on the complete space $X \times \mathbb{R}_+$ is Cauchy, we denote its limit by m .

Consider for $n, \ell \in \mathbb{N}$ also the probability measure $m_n^{\varepsilon_\ell}$ on $X \times \mathbb{R}_+$, given by

$$m_n^{\varepsilon_\ell} = \sum_{i \in M^{\varepsilon_\ell}} |B^{\varepsilon_\ell}(i)|_n \delta_{(i, v(i))} \quad \text{a. s.}$$

As there exists a. s. a coupling K' of the probability measures $m_n^{\varepsilon_\ell}$ and m^{ε_ℓ} with

$$K'\{(y, y)\} = m_n^{\varepsilon_\ell}\{y\} \wedge m^{\varepsilon_\ell}\{y\}$$

for all $y \in X \times \mathbb{R}_+$, the coupling characterization of the Prohorov metric (11.1) implies for each $k \in \mathbb{N}$

$$\begin{aligned} & d_P^{X \times \mathbb{R}_+}(m_n^{\varepsilon_\ell}, m^{\varepsilon_\ell}) \\ & \leq K'\{(y, y') \in (X \times \mathbb{R}_+)^2 : y \neq y'\} \\ & \leq m^{\varepsilon_\ell}\{(j, v(j)) : j \in M^{\varepsilon_\ell}, j > k\} + K'\{((i, v(i)), (j, v(j))) : i, j \in M^{\varepsilon_\ell}, i \neq j, j \leq k\} \\ & = \sum_{\substack{j \in M^{\varepsilon_\ell} \\ j > k}} |B^{\varepsilon_\ell}(j)| + \sum_{\substack{j \in M^{\varepsilon_\ell} \\ j \leq k}} ||B^{\varepsilon_\ell}(j)|_n - |B^{\varepsilon_\ell}(j)|| \quad \text{a. s.} \end{aligned}$$

Letting first n and then k tend to infinity, we deduce

$$\lim_{n \rightarrow \infty} d_P^{X \times \mathbb{R}_+}(m_n^{\varepsilon_\ell}, m^{\varepsilon_\ell}) = 0 \quad \text{a. s.} \quad (11.7)$$

Moreover, we define for each $n \in \mathbb{N}$ the probability measure

$$m_n = n^{-1} \sum_{i=1}^n \delta_{(i, v(i))}$$

on $X \times \mathbb{R}_+$. (Its first component corresponds to the measure on the starting points of the external branches that end in one of the first n leaves of T such that leaves are weighted according to the multiplicity given by the semi-metric ρ .) By (11.6),

$$d_P^{X \times \mathbb{R}_+}(m_n^{\varepsilon_\ell}, m_n) \leq 4\varepsilon_\ell \quad \text{a.s.} \quad (11.8)$$

for all $n, \ell \in \mathbb{N}$. From (11.6), (11.7), and (11.8), we obtain

$$m = \text{w-} \lim_{n \rightarrow \infty} m_n \quad \text{a.s.} \quad (11.9)$$

This shows the assertion. \square

In the first proof of Proposition 3.16, we constructed the sampling measure “by hand”. The idea for the second proof is to consider, in the metric completion of the coalescent tree associated with ρ , the closure of the subspace of the starting vertices of the external branches that end in the leaves labeled by the odd integers, and to show that this complete subspace contains the sequence of the starting vertices of the external branches associated with the even integers. To this aim, we use (11.4) from the first proof. The de Finetti theorem then yields the sampling measure for this exchangeable sequence. For a related result, see also Forman, Haulk, and Pitman [22], where trees are embedded into ℓ_1 .

Remark 11.2. The second proof given below goes in a direction that is similar to the argument in Section 7 of [21] for the construction of the sampling measure μ on the real tree $\mathbf{S} = \Gamma(\mathbf{T})$. That the equality $\Gamma(\mathbf{T}) = \Gamma(\mathbf{T}^-) = \Gamma(\mathbf{T}^+)$ on p. 40 in [21] holds for the embedding of $\Gamma(\mathbf{T}^-)$ and $\Gamma(\mathbf{T}^+)$ into $\Gamma(\mathbf{T})$ can be seen as in the proof below as $\Gamma(\mathbf{T})$, $\Gamma(\mathbf{T}^-)$, and $\Gamma(\mathbf{T}^+)$ then correspond to X , X_1 , and X_2 therein. The real tree $\Gamma(\mathbf{T}^-)$ can then be endowed with a measure like X_1 is endowed with μ^1 . Note that the starting vertices of the external branches and the subtree spanned by them are called the points of attachment and the core, respectively, in [21]. Another variant of the construction of the sampling measure is given in Remark 11.3 below.

We remark that the second last paragraph of the proof below shows that the isomorphism class of the weighted real tree (\mathbf{S}, μ) is a.s. equal to $\psi(r)$ where $(r, v) = \beta(d)$ and d is the exchangeable ultrametric on \mathbb{N} from [21, Section 7], which corresponds to ρ below. This equality can also be deduced from Theorem 3.18, Remark 3.7, as $\psi(r)$ is a.s. constant by the ergodicity assumption in [21], and from the Gromov reconstruction theorem.

Second proof of Proposition 3.16. Let $(r, v) = \beta(\rho)$. We construct the first component of the sampling measure, showing $r \in \mathfrak{D}^*$ a.s.

We denote by \mathbb{N}_1 the odd, and by \mathbb{N}_2 the even integers. Let (X, r) denote the metric completion of (\mathbb{N}, r) . A.s. by (11.3) and (11.4), there exists for each $i \in \mathbb{N}_2$ an integer $j \in \mathbb{N}_1$ with $r(i, j) < 2\varepsilon$. As ε can be chosen arbitrarily small, it follows that i is a.s. contained in the closure X_1 of the subset \mathbb{N}_1 of (X, r) a.s., hence $X_1 = X$ a.s. (Recall from Remark 2.2 that \mathbb{N} corresponds here to the set of starting vertices of the external branches in the coalescent tree (T, d) associated with ρ .)

For $i \in \mathbb{N}_1$, let

$$v^1(i) = \frac{1}{2} \inf_{j \in \mathbb{N}_1 \setminus \{i\}} \rho(i, j).$$

(This is the length of the external branch that ends in the leaf $(0, i)$ in the subtree spanned by the leaves with labels in \mathbb{N}_1 .) By exchangeability (hence contractability) of ρ and by definition of $v = \Upsilon(\rho)$, it follows that $v^1(i) = v(i)$ a.s. Let $\rho^1 = (\rho(i, j))_{i, j \in \mathbb{N}_1}$ be the restriction of ρ to \mathbb{N}_1 . We define the random variable $r^1 = (r^1(i, j))_{i, j \in \mathbb{N}_1}$ by

$$r^1(i, j) = (\rho^1(i, j) - v^1(i) - v^1(j)) \mathbf{1}\{i \neq j\}.$$

By definition of r in Section 2, it follows that $r^1 = (r(i, j))_{i, j \in \mathbb{N}_1}$ a.s.

Let Λ be a regular conditional distribution of ρ given ρ^1 . Then for a.a. ρ^1 , under $\Lambda(\rho^1, \cdot)$, the complete and separable metric space (X_1, r) is a.s. constant as r^1 is ρ^1 -measurable.

Moreover, the sequence $2, 4, 6, \dots$ of the even integers, viewed as a sequence in (X_1, r) , is exchangeable under $\Lambda(\rho^1, \cdot)$ for a.a. ρ^1 . To see this, we use that the Borel sigma algebra on (X_1, r) is generated by the balls around the elements of $\mathbb{N}_1 \subset X_1$. Let $n \in \mathbb{N}$, and let B_2, \dots, B_{2n} be some finite intersections of such balls. Note that $\{2 \in B_2, \dots, 2n \in B_{2n}\}$ can be written as an intersection of events of the form $\{\rho(i, j) < c\}$, where $i \in \mathbb{N}_2, j \in \mathbb{N}_1$ and $c \in (0, \infty)$. Using this, the uniqueness lemma, and the elementary fact that the conditional distribution of ρ given its restriction ρ^1 is invariant under permutations that leave \mathbb{N}_1 fixed, we obtain the claimed exchangeability.

For this exchangeable sequence, the de Finetti theorem yields, $\Lambda(\rho^1, \cdot)$ -a.s. for a.a. ρ^1 , a sampling measure μ^1 on (X_1, r) that is the weak limit of the probability measures $\mu_n^1 := n^{-1} \sum_{i=1}^n \delta_{2i}$ on (X_1, r) . By the same argument as above, also the closure X_2 of the subset \mathbb{N}_2 in (X, r) equals X a.s. On the event of probability 1 on which \mathbb{N}_2 is a dense subset of $X_2 = X = X_1$, an isometry $\varphi : X_1 \rightarrow X_2$ is given by $\varphi(i) = i$ for $i \in \mathbb{N}_2$. As also the weak limit of the image measures $\varphi(\mu_n^1)$ on (X_2, r) exists a.s., we have shown $(r(2i, 2j))_{i, j \in \mathbb{N}} \in \mathfrak{D}^*$ a.s. This implies $r \in \mathfrak{D}^*$ a.s. as r and $(r(2i, 2j))_{i, j \in \mathbb{N}}$ are equal in distribution by exchangeability of r .

That $(r, v) \in \hat{\mathfrak{D}}^*$ can be shown analogously by considering the sequence $(i, v(i))_{i \in \mathbb{N}_2}$ in the space $X_1 \times \mathbb{R}_+$ which we endow with the metric $d^{X_1 \times \mathbb{R}_+}((x', v'), (x'', v'')) = r(x', x'') \vee |v' - v''|$. \square

In the third proof, we use (11.4) to obtain on an event of probability 1 a labeling of the elements of the completion of (\mathbb{N}, r) that is determined by the equivalence class $\llbracket r \rrbracket$ of r with respect to finite permutations.

Remark 11.3. The sampling measure μ on the real tree $\mathbf{S} = \Gamma(\mathbf{T})$ from Remark 11.2 can also be obtained as in the third proof given below. Indeed, $i \in \mathbb{N}$ corresponds to the point of attachment $\Pi(i)$, and (Y, r) to the closure of $\{\Pi(i) : i \in \mathbb{N}\}$ in $\Gamma(\mathbf{T})$ in [21, Section 7]. When we label the elements of the closure of $\{\Pi(i) : i \in \mathbb{N}\}$ as in the proof below, we obtain a complete and separable metric space Y' that is a.s. constant by the ergodicity assumption in [21]. The de Finetti theorem, applied to the exchangeable sequence $\Pi(1), \Pi(2), \dots$ in Y' , then yields the sampling measure μ .

Third proof of Proposition 3.16. Let $(r, v) = \beta(\rho)$. We say that sequences $(y_j^1)_{j \in \mathbb{N}}$ and $(y_j^2)_{j \in \mathbb{N}}$ in (\mathbb{N}, r) are equivalent if $\lim_{j \rightarrow \infty} r(y_j^1, y_j^2) = 0$. Let Y be the set of the equivalence

classes of those Cauchy sequences $(y_j)_{j \in \mathbb{N}}$ in (\mathbb{N}, r) that satisfy $y_{j+1} > y_j$ for all $j \in \mathbb{N}$. We endow Y with the metric induced by r , namely $r(y^1, y^2) = \lim_{j \rightarrow \infty} r(y_j^1, y_j^2)$ for $y^1, y^2 \in Y$ with any representatives $(y_j^1)_{j \in \mathbb{N}}$ and $(y_j^2)_{j \in \mathbb{N}}$.

By (11.3) and (11.4), exchangeability of \sim^ε therein, and Kingman's correspondence, there exists a.s. for each $i \in \mathbb{N}$ and $\varepsilon > 0$ an arbitrarily large integer j with $r(i, j) < \varepsilon$. For each Cauchy sequence $(y_j^1)_{j \in \mathbb{N}}$ in (\mathbb{N}, r) , we thus find an equivalent Cauchy sequence $(y_j^2)_{j \in \mathbb{N}}$ in (\mathbb{N}, r) with $r(y_j^1, y_j^2) < 1/j$ and $y_{j+1}^2 > y_j^2$ for all $j \in \mathbb{N}$. Hence, we can identify each element i of \mathbb{N} with the equivalence class (of Cauchy sequences) in Y whose representatives are equivalent to the Cauchy sequence (i, i, \dots) in (\mathbb{N}, r) . Then (Y, r) is a metric completion of (\mathbb{N}, r) .

Let $\llbracket r \rrbracket$ denote the equivalence class (of distance matrices) of r with respect to finite permutations, and let Λ be a regular conditional distribution of r given $\llbracket r \rrbracket$. Note that finite permutations of r preserve the equivalence classes that are the elements of Y , and also the distances between these elements of Y . That is, the complete and separable metric space (Y, r) is determined by $\llbracket r \rrbracket$, hence a.s. constant under $\Lambda(\llbracket r \rrbracket, \cdot)$ for a.a. $\llbracket r \rrbracket$. We claim that for a.a. $\llbracket r \rrbracket$ under $\Lambda(\llbracket r \rrbracket, \cdot)$, the sequence $1, 2, 3, \dots$ is exchangeable in (Y, r) . Then, the de Finetti theorem yields the a.s. existence of the weak limit of the probability measures $n^{-1} \sum_{i=1}^n \delta_i$ on (Y, r) , which means that $r \in \mathfrak{D}^*$ a.s. Analogously, it can be seen that $(r, v) \in \mathfrak{D}^*$ a.s. by considering the sequence $(i, v(i))_{i \in \mathbb{N}}$ in the space $Y \times \mathbb{R}_+$ and a regular conditional distribution of (r, v) given the equivalence class of (r, v) with respect to finite permutations.

Now we carry out the proof of the claimed exchangeability of $1, 2, 3, \dots$. We fix $\llbracket r \rrbracket$ and hence (Y, r) . Let $n, \ell \in \mathbb{N}$. For $i \in [n]$, $k \in [\ell]$, let $y^{i,k}$ be an element of Y . Let the Cauchy sequences $(y_j^{i,k})_{j \in \mathbb{N}}$ in (\mathbb{N}, r) be representatives of the $y^{i,k}$. Here we can assume w.l.o.g. that $y_j^{i,k} > n$ for all $j \in \mathbb{N}$. Note that each $y^{i,k}$ has also the representative $(y_j^{i,k})_{j \in \mathbb{N}}$ in any (\mathbb{N}, r') with $\llbracket r' \rrbracket = \llbracket r \rrbracket$. Hence, as we fix $\llbracket r \rrbracket$, we can also fix the representatives $(y_j^{i,k})_{j \in \mathbb{N}}$ of the $y^{i,k}$. Let $(f^{i,k})$ be a collection of bounded continuous functions, and let p be a permutation of $[n]$. By exchangeability of r ,

$$\int \Lambda(\llbracket r \rrbracket, dr') \prod_{i \in [n], k \in [\ell]} f^{i,k} \circ r'(i, y_j^{i,k}) = \int \Lambda(\llbracket r \rrbracket, dr') \prod_{i \in [n], k \in [\ell]} f^{i,k} \circ r'(p(i), y_j^{i,k}) \quad (11.10)$$

for a.a. $\llbracket r \rrbracket$ and $j \in \mathbb{N}$. As in the first part of the proof, we can a.s. associate with each $i \in [n]$ a Cauchy sequence $(z_j^i)_{j \in \mathbb{N}}$ in (\mathbb{N}, r') that satisfies $\lim_{j \rightarrow \infty} r'(i, z_j^i) = 0$ and $z_{j+1}^i > z_j^i$ for all $j \in \mathbb{N}$. Then,

$$\lim_{j \rightarrow \infty} r'(i, y_j^{i,k}) = \lim_{j \rightarrow \infty} r'(z_j^i, y_j^{i,k}),$$

and similarly for the integrand on the right-hand side of equation (11.10). As $i \in [n]$ is identified with the equivalence class of $(z_j^i)_{j \in \mathbb{N}}$ in Y , and as each semi-metric r' on \mathbb{N} with $\llbracket r' \rrbracket = \llbracket r \rrbracket$ induces the same metric $r' = r$ on Y , letting j tend to infinity on both sides of equation (11.10) yields

$$\int \Lambda(\llbracket r \rrbracket, dr') \prod_{i \in [n], k \in [\ell]} f^{i,k} \circ r'(i, y^{i,k}) = \int \Lambda(\llbracket r \rrbracket, dr') \prod_{i \in [n], k \in [\ell]} f^{i,k} \circ r'(p(i), y^{i,k})$$

by dominated convergence. By approximating indicator variables with the $f^{i,k}$, we deduce that

$$\int \Lambda(\llbracket r \rrbracket, dr') \prod_{i \in [n], k \in [\ell]} \mathbf{1}\{r'(i, y^{i,k}) < c_{i,k}\} = \int \Lambda(\llbracket r \rrbracket, dr') \prod_{i \in [n], k \in [\ell]} \mathbf{1}\{r'(p(i), y^{i,k}) < c_{i,k}\}$$

for any numbers $c_{i,k} > 0$. The assertion follows by the uniqueness lemma as the balls of the form $\{y \in Y : r(y, y^{i,k}) < c_{i,k}\}$ generate the Borel sigma algebra on (Y, r) . \square

12 Construction from the flow of bridges

A random non-decreasing right-continuous function $\tilde{F} : [0, 1] \rightarrow [0, 1]$ with exchangeable increments and $\tilde{F}(0) = 0$, $\tilde{F}(1) = 1$ is called a bridge. We view a bridge as a random variable with values in the space of càdlàg paths $[0, 1] \rightarrow [0, 1]$ which we endow with the Skorohod metric. The dual flow of bridges of Bertoin and Le Gall [5] is a collection $F = (F_{s,t}, s < t)$ of bridges that satisfies the following properties (see [5, Section 5.1]):

- (i) For every $s < t < u$, $F_{t,u} \circ F_{s,t} = F_{s,u}$ a.s.
- (ii) The law of $F_{s,t}$ depends only on $t - s$. For $s_1 < s_2 < \dots < s_n$, the bridges $F_{s_1, s_2}, F_{s_2, s_3}, \dots, F_{s_{n-1}, s_n}$ are independent.
- (iii) $F_{0,0}$ is the identity function. For every $x \in [0, 1]$, the random variable $F_{0,t}(x)$ converges to x in probability as t decreases to zero.

For each $s < t$, it is also assumed that $F_{s,t}$ is a.s. not the identity function.

The interpretation is that the individuals of a continuous population are represented by the elements of the interval $[0, 1]$. For each $s \leq t$, the individuals in a subinterval $(x_1, x_2]$ at time s have descendants at time t that are a.s. the elements of $(F_{s,t}(x_1), F_{s,t}(x_2)]$, see [7].

In [5, Section 3], Kingman's correspondence is extended so as to represent distributions of Ξ -coalescents in terms of sampling from flows of bridges. Let F be a dual flow of bridges, and let $V = (V_i, i \in \mathbb{N})$ be an iid sequence of uniform $[0, 1]$ -valued random variables, independent of F . This iid sequence is interpreted as a sequence of random samples from the population at some time $t \in \mathbb{R}$. For each $s \in \mathbb{R}_+$, a partition $\tilde{\pi}_s^{(t)}$ is defined such that any integers $i, j \in \mathbb{N}$ are in the same block of $\tilde{\pi}_s^{(t)}$ if and only if $F_{t-s,t}^{-1}(V_i) = F_{t-s,t}^{-1}(V_j)$ which means that these samples have the same ancestor at time $t - s$. Here we set $f^{-1}(t) = \inf\{s \in [0, 1] : f(s) > t \text{ or } s = 1\}$ for $t \in [0, 1]$ and a càdlàg function $f : [0, 1] \rightarrow [0, 1]$. In [5, Theorem 1], it is shown that the partition-valued process $(\tilde{\pi}_s^{(t)}, s \in \mathbb{R}_+)$ obtained in this way is a version of a Ξ -coalescent of Schweinsberg [44].

For each $t \in \mathbb{R}$, there exists an event of probability 1 on which for all $s \leq s' \in \mathbb{Q}_+$, the partition $\tilde{\pi}_{s'}^{(t)}$ can be obtained by merging blocks of the partition $\tilde{\pi}_s^{(t)}$. We can thus define a.s. an ultrametric $\tilde{\rho}_t$ by

$$\tilde{\rho}_t(i, j) = 2 \inf\{s \in \mathbb{Q}_+ : i \text{ and } j \text{ are in the same block of } \tilde{\pi}_s^{(t)}\}.$$

The assumption that for each $r < s$, the bridge $F_{r,s}$ is a.s. not the identity function implies that the infimum in the definition of $\tilde{\rho}_t(i, j)$ is a.s. not over the empty set.

Moreover, we define a.s. a random variable $\tilde{\nu}_t$ with values in the space (\mathcal{U}, d_P) of exchangeable distributions on \mathfrak{U} such that $\tilde{\nu}_t$ is a regular conditional distribution of $\tilde{\rho}_t$ given the collection of bridges $(F_{t-s,t}, s \in \mathbb{Q}_+)$. For the existence of this regular conditional distribution, see e.g. [28, Theorem 6.3].

Analogously to Sections 8.2 and 8.3, for a finite measure Ξ on Δ , a stationary \mathcal{U}^{erg} -valued Ξ -Fleming-Viot process $(\nu_t, t \in \mathbb{R})$ is given by $\nu_t = \alpha(\nu^{\hat{\psi} \circ \beta(\bar{\rho}_t)})$, where $(\bar{\rho}_t, t \in \mathbb{R})$ is defined as in Section 10. We note that a stationary \mathcal{U}^{erg} -valued Ξ -Fleming-Viot process can be read off from the dual flow of bridges:

Theorem 12.1. *There exists a finite measure Ξ on Δ such that the process $(\tilde{\nu}_t, t \in \mathbb{R})$ is a version of a stationary \mathcal{U}^{erg} -valued Ξ -Fleming-Viot process.*

For the proof of Theorem 12.1, we show that $(\tilde{\nu}_t, t \in \mathbb{R})$ is a Markov process and has the transition kernel of a \mathcal{U}^{erg} -valued Ξ -Fleming-Viot process. In the following, we fix $u \in \mathbb{R}_+$.

First, we define for each finite measure Ξ on Δ a probability kernel Λ_Ξ from \mathcal{U} to \mathfrak{U} such that for each $\nu \in \mathcal{U}$, the distribution $\Lambda_\Xi(\nu, \cdot)$ is the distribution of a random variable ρ which we define as follows. Let ρ' be a random variable with distribution ν . Let ρ'' be an independent \mathfrak{U} -valued random variable that is distributed as the random ultrametric associated with a Ξ -coalescent. That is, ρ'' shall be distributed as the random variable $\bar{\rho}_u$ mentioned above, cf. Remark 6.4. We define a partition π of \mathbb{N} such that i and j are in the same block of π if and only if $\rho''(i, j) < 2u$. Let $B_1(\pi), B_2(\pi), \dots$ be the blocks of π , ordered increasingly according to their smallest element. For $i \in \mathbb{N}$, let $A(i)$ be the integer j such that $i \in B_j(\pi)$.

Then we set for $i, j \in \mathbb{N}$

$$\rho(i, j) = \begin{cases} \rho''(i, j) \wedge (2u) & \text{if } \rho''(i, j) < 2u \\ 2u + \rho'(A(i), A(j)) & \text{else.} \end{cases}$$

In the following, we also fix $t \in \mathbb{R}$.

Remark 12.2. Let $(\bar{\rho}_s, s \in \mathbb{R})$ be defined as in Section 10 from a measure Ξ , and let $\nu_t = \alpha(\nu^{\hat{\psi} \circ \beta(\bar{\rho}_t)})$. Note that Λ_Ξ is a regular conditional distribution of $\bar{\rho}_{t+u}$ given ν_t . This follows as $\bar{\rho}_{t+u} \wedge (2u)$ is independent of ν_t and $\bar{\rho}_t$, as $\bar{\rho}_{t+u}(i, j) = 2u + \bar{\rho}_t(A_t(t+u, i), A_t(t+u, j))$ for $i, j \in \mathbb{N}$ with $\bar{\rho}_{t+u}(i, j) \geq 2u$, as $A_t(t+u, i)$ can be read off from $\bar{\rho}_{t+u} \wedge (2u)$ like $A(i)$ can be read off from ρ'' in the definition of Λ_Ξ , and as ν_t is a regular conditional distribution of $\bar{\rho}_t$ given ν_t by Remark 3.19.

Lemma 12.3. *There exists a finite measure Ξ on Δ such that Λ_Ξ is a regular conditional distribution of $\tilde{\rho}_{t+u}$ given $\tilde{\nu}_t$. Moreover, $\tilde{\rho}_{t+u}$ is conditionally independent of $(\tilde{\nu}_s, s \leq t)$ given $\tilde{\nu}_t$.*

Proof. Let $k \in \mathbb{N}$ and $0 = s_1 \leq s_2 \leq \dots \leq s_k$. We claim that given the collection of bridges $(F_{r,s} : r < s \leq t)$, the random variable $\tilde{\rho}_{t+u}$ has conditional distribution $\Lambda_\Xi(\tilde{\nu}_t, \cdot)$ for some finite measure Ξ on Δ . By construction of $\tilde{\nu}_s$, this claim implies both assertions of the lemma.

We assume that the coalescent process $(\tilde{\pi}_s^{(t+u)}, s \in \mathbb{R}_+)$ and the associated ultrametric $\tilde{\rho}_{t+u}$ are constructed as above from F and a sequence $(V_i, i \in \mathbb{N})$ of independent uniformly

distributed $[0, 1]$ -valued random variables that is independent of F . By [5, Theorem 1], there exists a finite measure Ξ on Δ such that $(\tilde{\pi}_s^{(t+u)}, s \in \mathbb{R}_+)$ is a version of a Ξ -coalescent.

Let $B_1(\tilde{\pi}_u^{(t+u)}), B_2(\tilde{\pi}_u^{(t+u)}), \dots$ be the blocks of $\tilde{\pi}_u^{(t+u)}$ in increasing order according to their respective smallest element. For $i \in \mathbb{N}$, we define $\tilde{A}(i) = j$ where j is the integer such that $i \in B_j(\tilde{\pi}_u^{(t+u)})$. We define a sequence $V' = (V'_i, i \in \mathbb{N})$ analogously to equation (3) of [5]: For $i \in \mathbb{N}$ with $i \leq \#\tilde{\pi}_u^{(t+u)}$, we set $V'_i = F_{t,t+u}^{-1}(V_j)$, where j is any element of $B_i(\tilde{\pi}_u^{(t+u)})$. If the number of blocks $\#\tilde{\pi}_u^{(t+u)}$ is finite, we extend the sequence $(V'_i, i \leq \#\tilde{\pi}_u^{(t+u)})$ to $(V'_i, i \in \mathbb{N})$ using an independent sequence of independent uniform random variables on $[0, 1]$.

Let $\ell \in \mathbb{N}$ and $0 \leq u_1 \leq u_2 \leq \dots \leq u_\ell = u$. Repeated application of [5, Lemma 2] to the bridges $F_{t+u-u_1, t+u}, \dots, F_{t+u-u_\ell, t+u-u_{\ell-1}}$ (similarly to [5, Corollary 1]) yields that V' is a sequence of independent $[0, 1]$ -valued uniformly distributed random variables that is also independent of $\tilde{\pi}_{u_1}^{(t+u)}, \dots, \tilde{\pi}_{u_\ell}^{(t+u)}$. By construction and property (ii) of the dual flow of bridges, V' and $\tilde{\pi}_{u_1}^{(t+u)}, \dots, \tilde{\pi}_{u_\ell}^{(t+u)}$ are also independent of $(F_{r,s} : r < s \leq t)$.

We define $\tilde{\rho}_t$ from F and the sequence V' . Then $\tilde{\rho}_t$ is conditionally independent of $\tilde{\rho}_{t+u} \wedge (2u)$ given the collection of bridges $(F_{r,s} : r < s \leq t)$. This follows from the above by the uniqueness lemma as for $i, j \in \mathbb{N}$ and $m = 1, \dots, \ell$, $\{\tilde{\rho}_{t+u}(i, j) \leq u_m\}$ is, up to null events, the event that i and j are in the same block of $\tilde{\pi}_{u_m}^{(t+u)}$.

By construction, $\tilde{\nu}_t$ is a conditional distribution of $\tilde{\rho}_t$ given $(F_{r,s} : r < s \leq t)$. We also define the coalescent process $(\tilde{\pi}_s^{(t)}, s \in \mathbb{R}_+)$ from V' and F . Then $\tilde{\rho}_t$ is the associated ultrametric. For $i, j \in \mathbb{N}$ and $s \in \mathbb{R}_+$, the following events are equal up to null events:

$$\begin{aligned} \{\tilde{\rho}_{t+u}(i, j) \leq 2(u+s)\} &= \{i, j \text{ are in the same block of } \tilde{\pi}_{u+s}^{(t+u)}\} \\ &= \{\tilde{A}(i), \tilde{A}(j) \text{ are in the same block of } \tilde{\pi}_s^{(t)}\} = \{\tilde{\rho}_t(\tilde{A}(i), \tilde{A}(j)) \leq 2s\}. \end{aligned}$$

For the equality up to null events of the second and the third event, we use the definition of V' and property (i) of the dual flow of bridges. It follows that a.s.,

$$\tilde{\rho}_{t+u}(i, j) = \begin{cases} \tilde{\rho}_{t+u}(i, j) \wedge (2u) & \text{if } \tilde{\rho}_{t+u}(i, j) < 2u \\ 2u + \tilde{\rho}_t(\tilde{A}(i), \tilde{A}(j)) & \text{else.} \end{cases}$$

The claim follows as $\tilde{A}(i)$ can a.s. be read off from $\tilde{\rho}_{t+u} \wedge (2u)$ in the same way as $A(i)$ is read off from ρ'' in the definition of Λ_Ξ . \square

To deduce Theorem 12.1, we use that $\tilde{\nu}_{t+u} \in \mathcal{U}^{\text{erg}}$ a.s.

Proof of Theorem 12.1. Let $t \in \mathbb{R}$ and $u > 0$. By Remark 3.22 and as the sequence V in the definition of $\tilde{\rho}_t$ is iid, $\tilde{\nu}_t \in \mathcal{U}^{\text{erg}}$ a.s. This can be seen directly or by an application of e.g. [29, Lemma 7.35]. By construction, $\tilde{\nu}_{t+u}$ is a regular conditional distribution of $\tilde{\rho}_{t+u}$ given $\tilde{\nu}_{t+u}$. By Corollary 3.24, $\tilde{\nu}_{t+u} = \zeta(\tilde{\rho}_{t+u})$ a.s., where $\zeta : \mathfrak{U} \rightarrow \mathcal{U}^{\text{erg}}, \rho \mapsto \nu^{\hat{\psi} \circ \beta(\rho)}$. Hence, by Lemma 12.3, there exists a finite measure Ξ on Δ such that $\Lambda_\Xi(\cdot, \zeta^{-1}(\cdot))$ is a regular conditional distribution of $\tilde{\nu}_{t+u}$ given $\tilde{\nu}_t$, and $\tilde{\nu}_{t+u}$ is conditionally independent of $(\tilde{\nu}_s, s \leq t)$ given $\tilde{\nu}_t$. The latter property is the Markov property of $(\tilde{\nu}_s, s \in \mathbb{R})$.

Let now $(\nu_s, s \in \mathbb{R})$ be a \mathcal{U}^{erg} -valued Ξ -Fleming-Viot process defined from $(\tilde{\rho}_s, s \in \mathbb{R})$ as recalled in the beginning of this section. As in Theorem 4.2 (or alternatively, by an

extension of Remark 12.2), the process $(\nu_t, t \in \mathbb{R})$ is Markovian. By Remark 12.2 and as $\nu_{t+u} = \zeta(\bar{\rho}_{t+u})$ by definition, $\Lambda_{\Xi}(\cdot, \zeta^{-1}(\cdot))$ is a regular conditional distribution also of ν_{t+u} given ν_t . This implies the assertion. \square

List of notation

Here we collect notation that is used globally in the article.

Miscellaneous

$\mathbb{R}_+ = [0, \infty)$, $\mathbb{Q}_+ = \mathbb{R}_+ \cap \mathbb{Q}$, $\mathbb{N} = \{1, 2, 3, \dots\}$, $[n] = \{1, \dots, n\}$ for $n \in \mathbb{N}$, $[0] = \emptyset$
 γ_n : restriction map in various contexts, (p. 12/l.-8, p. 16, p. 19/l. 19)

(Marked) distance matrices

\mathfrak{U} : space of semi-ultrametrics on \mathbb{N} , (p. 7/l. 14)
 $\hat{\mathfrak{U}}$: space of decomposed semi-ultrametrics on \mathbb{N} , (p. 7/l. 21)
 $\mathfrak{D}, \hat{\mathfrak{D}}$: spaces of (decomposed) semimetrics on \mathbb{N} (p. 7/l. 14, p. 34/l. -15)
 α : map that retrieves the semi-ultrametric from a decomposed semi-ultrametric (p. 7/l. 19)
 $\beta : \mathfrak{U} \rightarrow \hat{\mathfrak{U}}$: decomposition map into the external branches and the remaining subtree (p. 7/l. -12)
 $\Upsilon(\rho)$: vector of the lengths of the external branches in the coalescent tree associated with ρ (p. 7/l. -13)

(Marked) metric measure spaces

\mathbb{M} : space of isomorphy classes of metric measure spaces (p. 8/l. 14)
 \mathbb{U} : space of isomorphy classes of ultrametric measure spaces (p. 8/l. -1)
 $\hat{\mathbb{M}}, \hat{\mathbb{U}}$: spaces of isomorphy classes of marked metric measure spaces (p. 8/l. 18, p. 9/l. 3)
 ν^χ : distance matrix distribution of $\chi \in \mathbb{M}$ (p. 3/l. -10, p. 8/l. 19) or marked distance matrix distribution of $\chi \in \hat{\mathbb{M}}$ (p. 4/l. 9, p. 8/l. 20)
 \mathcal{U}^{erg} : space of distance matrix distributions (p. 14/l. -1)
 $\psi : \mathfrak{D} \rightarrow \mathbb{M}$, $\hat{\psi} : \mathfrak{D} \times \mathbb{R}_+^{\mathbb{N}} \rightarrow \hat{\mathbb{M}}$: construction of (marked) metric measure spaces (p. 11/l. -9, p. 11/l. -3)
 $\mathfrak{D}^*, \hat{\mathfrak{D}}^*$: sets of (marked) distance matrices with a good sampling measure (p. 11/l. -5, p. 12/l. 2)
 $\mathcal{C}_n, \mathcal{C}, \hat{\mathcal{C}}_n, \hat{\mathcal{C}}$: sets of bounded differentiable functions with bounded uniformly continuous derivative (p. 16)
 Π : set of polynomials on \mathbb{U} (p. 16/l. -12)
 $\hat{\Pi}$: set of marked polynomials on $\hat{\mathbb{U}}$ (p. 16/l. -10)
 \mathcal{C} : a set of test functions on \mathcal{U}^{erg} (p. 16/l. -8)

Partitions and semi-partitions

\mathcal{P} : Set of partitions of \mathbb{N}
 $B_i(\pi)$: i -th block of a partition π (p. 19/l. -19)
 $\#\pi$: number of blocks of a partition π
 $K_{i,j}$: partition of \mathbb{N} that contains only $\{i, j\}$ and singleton blocks (p. 21/l. -10)
 \mathcal{P}_n : Set of partitions of $[n]$, associated transformations (equation (5.2))
 $\mathbf{0}_n = \{\{1\}, \dots, \{n\}\} \in \mathcal{P}_n$
 \mathcal{P}^n : Set of partitions of \mathbb{N} in which the first n integers are not all in different blocks (p. 19/l. 22)
 $\hat{\mathcal{P}}^n$: Set of partitions of \mathbb{N} in which the first n integers are not all in singleton blocks (p. 26/l. 4)
 \mathcal{S}_n set of semi-partitions of $[n]$, associated transformations (p. 25/l. -19, p. 25/l. -12)

$\Delta = \{x = (x_1, x_2, \dots) : x_1 \geq x_2 \geq \dots \geq 0, |x|_1 \leq 1\}$
 $\kappa(x, \cdot)$: paintbox distribution associated with $x \in \Delta$ (p. 21/l. 24)

Genealogy in the lockdown model

η : point measure on $(0, \infty) \times \mathcal{P}$ that encodes the reproduction events, (p. 19/l. -13, p. 21/l. -7)
 $A_s(t, i)$: level of the ancestor at time s of the particle on level i at time t (p. 19/l. -3)
 $\rho_t(i, j)$: genealogical distance (p. 20/l. 12)
 (r_t, v_t) : decomposed genealogical distance (p. 25)
 $\Xi = \Xi_0 + \Xi\{0\}\delta_0$, equation (6.1)
 H_Ξ : characteristic measure of η (p. 21/l. -8)

References

- [1] D. Aldous, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR1207226
- [2] T. Austin, *Exchangeable random measures*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 3, 842–861. MR3365963
- [3] J. Berestycki, N. Berestycki, and J. Schweinsberg, *Small-time behavior of beta coalescents*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 2, 214–238. MR2446321
- [4] J. Bertoin, *Random fragmentation and coagulation processes*, Cambridge University Press, Cambridge, 2006. MR2253162
- [5] J. Bertoin and J.-F. Le Gall, *Stochastic flows associated to coalescent processes*, Probab. Theory Related Fields **126** (2003), no. 2, 261–288. MR1990057
- [6] ———, *Stochastic flows associated to coalescent processes. II. Stochastic differential equations*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), no. 3, 307–333. MR2139022
- [7] ———, *Stochastic flows associated to coalescent processes. III. Limit theorems*, Illinois J. Math. **50** (2006), no. 1-4, 147–181 (electronic). MR2247827
- [8] M. Birkner, J. Blath, M. Möhle, M. Steinrücken, and J. Tams, *A modified lockdown construction for the Xi-Fleming-Viot process with mutation and populations with recurrent bottlenecks*, ALEA Lat. Am. J. Probab. Math. Stat. **6** (2009), 25–61. MR2485878
- [9] P. Carmona, F. Petit, and M. Yor, *Beta-gamma random variables and intertwining relations between certain Markov processes*, Rev. Mat. Iberoamericana **14** (1998), no. 2, 311–367. MR1654531
- [10] I. Dahmer and G. Kersting, *The total external length of the evolving Kingman coalescent*, Probab. Theory Related Fields (2016). <http://dx.doi.org/10.1007/s00440-016-0703-7>.
- [11] I. Dahmer, R. Knobloch, and A. Wakolbinger, *The Kingman tree length process has infinite quadratic variation*, Electron. Commun. Probab. **19** (2014), no. 87. MR3298276
- [12] A. Depperschmidt, A. Greven, and P. Pfaffelhuber, *Marked metric measure spaces*, Electron. Commun. Probab. **16** (2011), 174–188. MR2783338
- [13] ———, *Tree-valued Fleming-Viot dynamics with mutation and selection*, Ann. Appl. Probab. **22** (2012), no. 6, 2560–2615. MR3024977
- [14] ———, *Path-properties of the tree-valued Fleming-Viot process*, Electron. J. Probab. **18** (2013), no. 84. MR3109623
- [15] P. Donnelly and T. G. Kurtz, *A countable representation of the Fleming-Viot measure-valued diffusion*, Ann. Probab. **24** (1996), no. 2, 698–742. MR1404525
- [16] ———, *Particle representations for measure-valued population models*, Ann. Probab. **27** (1999), no. 1, 166–205. MR1681126

- [17] A. Etheridge and T. G. Kurtz, *Genealogical constructions of population models* (2016). arXiv:1402.6724.
- [18] S. N. Ethier and T. G. Kurtz, *Markov processes: Characterization and convergence*, Wiley, New York, 1986. MR838085
- [19] S. N. Evans, *Kingman's coalescent as a random metric space*, Stochastic models (Ottawa, ON, 1998), 2000, pp. 105–114. MR1765005
- [20] ———, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR2351587
- [21] S. N. Evans, R. Grübel, and A. Wakolbinger, *Doob–Martin boundary of Rémy's tree growth chain* (2016). arXiv:1411.2526. To appear in Ann. Prob.
- [22] N. Forman, C. Haulk, and J. Pitman, *A representation of exchangeable hierarchies by sampling from real trees* (2015). arXiv:1101.5619.
- [23] C. Foucart, *Generalized Fleming–Viot processes with immigration via stochastic flows of partitions*, ALEA Lat. Am. J. Probab. Math. Stat. **9** (2012), no. 2, 451–472. MR3069373
- [24] A. Greven, P. Pfaffelhuber, and A. Winter, *Convergence in distribution of random metric measure spaces (Λ -coalescent measure trees)*, Probab. Theory Related Fields **145** (2009), no. 1-2, 285–322. MR2520129
- [25] ———, *Tree-valued resampling dynamics Martingale problems and applications*, Probab. Theory Related Fields **155** (2013), no. 3-4, 789–838. MR3034793
- [26] S. Gufler, *Convergence of tree-valued Cannings chains* (2016). arXiv:1608.08203.
- [27] ———, *Pathwise construction of tree-valued Fleming–Viot processes* (2016). arXiv:1404.3682.
- [28] O. Kallenberg, *Foundations of modern probability*, Second Edition, Springer, New York, 2002. MR1876169
- [29] ———, *Probabilistic symmetries and invariance principles*, Springer, New York, 2005. MR2161313
- [30] G. Kersting, J. Schweinsberg, and A. Wakolbinger, *The evolving beta coalescent*, Electron. J. Probab. **19** (2014), no. 64. MR3238784
- [31] J. F. C. Kingman, *The coalescent*, Stochastic Process. Appl. **13** (1982), no. 3, 235–248. MR671034
- [32] S. Kliem and W. Löhner, *Existence of mark functions in marked metric measure spaces*, Electron. J. Probab. **20** (2015), no. 73. MR3371432
- [33] T. G. Kurtz, *Martingale problems for conditional distributions of Markov processes*, Electron. J. Probab. **3** (1998), no. 9, 29 pp. MR1637085
- [34] T. G. Kurtz and G. Nappo, *The filtered martingale problem*, The Oxford handbook of nonlinear filtering, 2011, pp. 129–165. MR2884595
- [35] C. Labbé, *From flows of Λ -Fleming–Viot processes to lockdown processes via flows of partitions*, Electron. J. Probab. **19** (2014), no. 55. MR3227064
- [36] W. Löhner, *Equivalence of Gromov–Prohorov- and Gromov's \square_λ -metric on the space of metric measure spaces*, Electron. Commun. Probab. **18** (2013), no. 17. MR3037215
- [37] M. Möhle, *Asymptotic results for coalescent processes without proper frequencies and applications to the two-parameter Poisson–Dirichlet coalescent*, Stochastic Process. Appl. **120** (2010), no. 11, 2159–2173. MR2684740
- [38] M. Möhle and S. Sagitov, *A classification of coalescent processes for haploid exchangeable population models*, Ann. Probab. **29** (2001), no. 4, 1547–1562. MR1880231
- [39] P. Pfaffelhuber and A. Wakolbinger, *The process of most recent common ancestors in an evolving coalescent*, Stochastic Process. Appl. **116** (2006), no. 12, 1836–1859. MR2307061

- [40] P. Pfaffelhuber, A. Wakolbinger, and H. Weisshaupt, *The tree length of an evolving coalescent*, Probab. Theory Related Fields **151** (2011), no. 3-4, 529–557. MR2851692
- [41] J. Pitman, *Coalescents with multiple collisions*, Ann. Probab. **27** (1999), no. 4, 1870–1902. MR1742892
- [42] L. C. G. Rogers and J. W. Pitman, *Markov functions*, Ann. Probab. **9** (1981), no. 4, 573–582. MR624684
- [43] S. Sagitov, *The general coalescent with asynchronous mergers of ancestral lines*, J. Appl. Probab. **36** (1999), no. 4, 1116–1125. MR1742154
- [44] J. Schweinsberg, *Coalescents with simultaneous multiple collisions*, Electron. J. Probab. **5** (2000), no. 4. MR1781024
- [45] V. S. Varadarajan, *Groups of automorphisms of Borel spaces*, Trans. Amer. Math. Soc. **109** (1963), 191–220. MR0159923
- [46] A. Véber and A. Wakolbinger, *The spatial Lambda-Fleming-Viot process: an event-based construction and a lookdown representation*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 570–598. MR3335017
- [47] A. M. Vershik, *Random and universal metric spaces*, Dynamics and randomness II, Kluwer Acad. Publ., Dordrecht, 2004, pp. 199–228. MR2153330

Acknowledgments. The author thanks Götz Kersting, Anton Wakolbinger, and two anonymous referees for helpful comments. Partial support from the DFG Priority Programme 1590 “Probabilistic Structures in Evolution” is acknowledged.